

## REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Service Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

### PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE (DD/MM/YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)			
08/29/2011	Final	June 2008 - May 2011			
4. TITLE AND SUBTITLE		5a. CONTRACT NUMBER			
(YIP 08) Structural Hierarchies in Biomimetic Materials: Protein Inspired De Novo Materials		FA9550-08-1-0321			
6. AUTHOR(S)		5b. GRANT NUMBER			
Buehler, Markus, J.		5c. PROGRAM ELEMENT NUMBER			
5d. PROJECT NUMBER		5e. TASK NUMBER			
5f. WORK UNIT NUMBER		5g. PERFORMING ORGANIZATION NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)		8. SPONSOR/MONITOR'S ACRONYM(S)			
Massachusetts Institute of Technology Office of Sponsored Research 77 Massachusetts Avenue E19-750 Cambridge, MA 02139		AFOSR			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S REPORT NUMBER(S)			
Air Force Office of Scientific Research (AFOSR) Dr. B. L. ("Les") Lee, Sc.D. Program Manager for Mechanics of Multifunctional Materials & Microsystems, 875 N. Randolph Street, AFOSR/NA, Suite 325, Room 3112, Arlington, VA 22203		AFRL-OSR-VA-TR-2012-0267			
12. DISTRIBUTION/AVAILABILITY STATEMENT		13. SUPPLEMENTARY NOTES			
Unlimited		N/A			
14. ABSTRACT		Hierarchical materials represent intriguing examples of multi-functional systems that combine disparate properties such as robustness, high strength, high elasticity changeability, controllability and the ability to self-assemble and self-heal. In this project we accomplished a multi-scale analysis of the fundamental material concepts that make it possible to achieve these properties. The work bridged the gap between biological and engineering sciences, facilitating the design of de novo biomimetic structures and materials with similar properties. We employed an innovative approach that combined theoretical analyses, large-scale atomistic based multi-scale simulation implemented on massively parallelized supercomputers with experimental work.			
15. SUBJECT TERMS		Mechanics, multiscale modeling, materialomics, intermediate filaments, proteins, protein materials, biologically inspired, biomimetic			
16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT c. THIS PAGE		17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Markus Buehler (mbuehler@MIT.EDU)	19b. TELEPHONE NUMBER (include area code) 617 452 2750
		UU	UU	<input type="button" value="Reset"/>	

Standard Form 298 (Rev. 8/98)  
Prescribed by ANSI Std. Z39-18

Adobe Professional 7.0

## **14. Abstract**

Hierarchical materials represent intriguing examples of multi-functional systems that combine disparate properties such as robustness, high strength, high elasticity, changeability, controllability and the ability to self-assemble and self-heal. In this project we accomplished a multi-scale analysis of the fundamental material concepts that make it possible to achieve these properties. The work bridged the gap between biological and engineering sciences, facilitating the design of *de novo* biomimetic structures and materials with similar properties. We employed an innovative approach that combined theoretical analyses, large-scale atomistic based multi-scale simulation implemented on massively parallelized supercomputers with experimental work. Our efforts focused on mechanical properties including elasticity, fracture resistance, self-assembly, and how these properties can be controlled. A focal point was the role and utilization of material hierarchies, an abundant trait of all protein materials and critical to their ability to combine disparate material properties. We focused on alpha-helical intermediate filament motif found in the cell's cytoskeleton, also forming the basis of wool and hair in the study of these scientific principles and covered a vast range of scales from nano to macro. Our work provided quantitative predictions of the elastic and strength properties of protein materials throughout vast range of time scales.



# YIP 08: Structural Hierarchies in Biomimetic Materials

Markus J. Buehler (Massachusetts Institute of Technology, [mbuehler@MIT.EDU](mailto:mbuehler@MIT.EDU))

Grant number FA9550-08-1-0321

## Multi-scale approach to understand biological protein materials & translate design concepts into engineering applications

### Current engineering materials

- Remain limited in their ability to combine disparate properties such as high strength, robustness, self-healing, mutability

### Biological materials and structures

- Show intriguing material properties • combine disparate properties in a single material

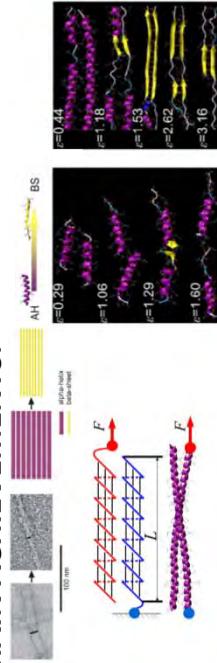
### De novo materials design

- Requires bottom-up structural design, from nano to macro • need to understand structure-property links in biological materials

### Multi-scale analysis of biological structures

- Intermediated filaments, found in cells nuclear envelop, provide intriguing mechanical properties • provide great extensibility, strength, mechanical robustness, and ability to self-heal • structural analysis reveals a intricate design, from atomistic to macroscopic

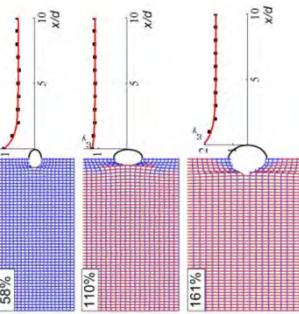
### MAIN ACHIEVEMENTS:



↑ Mechanism of the  $\alpha$ - $\beta$  transition caused by mechanical force.  
Oin, Krepak, Buehler, PRL, 2010, PLOS ONE, 2009

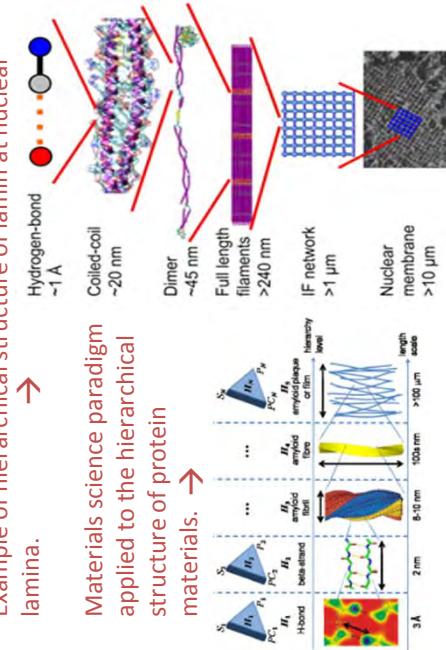
### Structural transition related

### Non-linear mechanical property



→ Study of nuclear lamina (meshwork of filaments).  
Discovered the flaw tolerance of this structure is caused by non-linear mechanical response of intermediate filaments.

### Example of hierarchical structure of laminin at nuclear lamina.



→ Materials science paradigm applied to the hierarchical structure of protein materials.

### HOW IT WORKS:

- The non-linear mechanical response of hierarchical structure of proteins is caused by the structural transition mechanisms during deformation.
- The alterable structure enables proteins to combine disparate material properties (flexible, strength, robustness).
- The structural and mechanical property of intermediated filaments are altered by point mutation.

### QUANTITATIVE IMPACT

### IMPACT

- Provide the first atomic mechanism and condition of the stiffening behavior of alpha-helical materials.
- Developed model to link the microscopic structural transition with the macroscopic behavior.
- Enables to seek the mechanical property of protein materials with mutation.

### TRANSITIONS

- New MURI project for fiber design • PI visited AFRL in 2010 (Wright-Patterson AFB) • PI awarded PECASE

### FUTURE RESEARCH GOALS

- Facilitate merger of structure, material property and function through the bottom-up multi-scale design, from atomic to macroscopic behavior.
- Quantitatively understand the assembling process of protein materials and factors to affect the material property of those assembled structures.
- Quantitatively understand the cascaded activation of mutations and their effect on material performance.
- Provide basis for new engineering paradigm to functionalize the hierarchical structural materials (e.g. new composites or polymers combine high flexibility, ultimate strength, self healing and robustness): novel materials for advanced Air Force technologies

### END-OF-PHASE GOAL

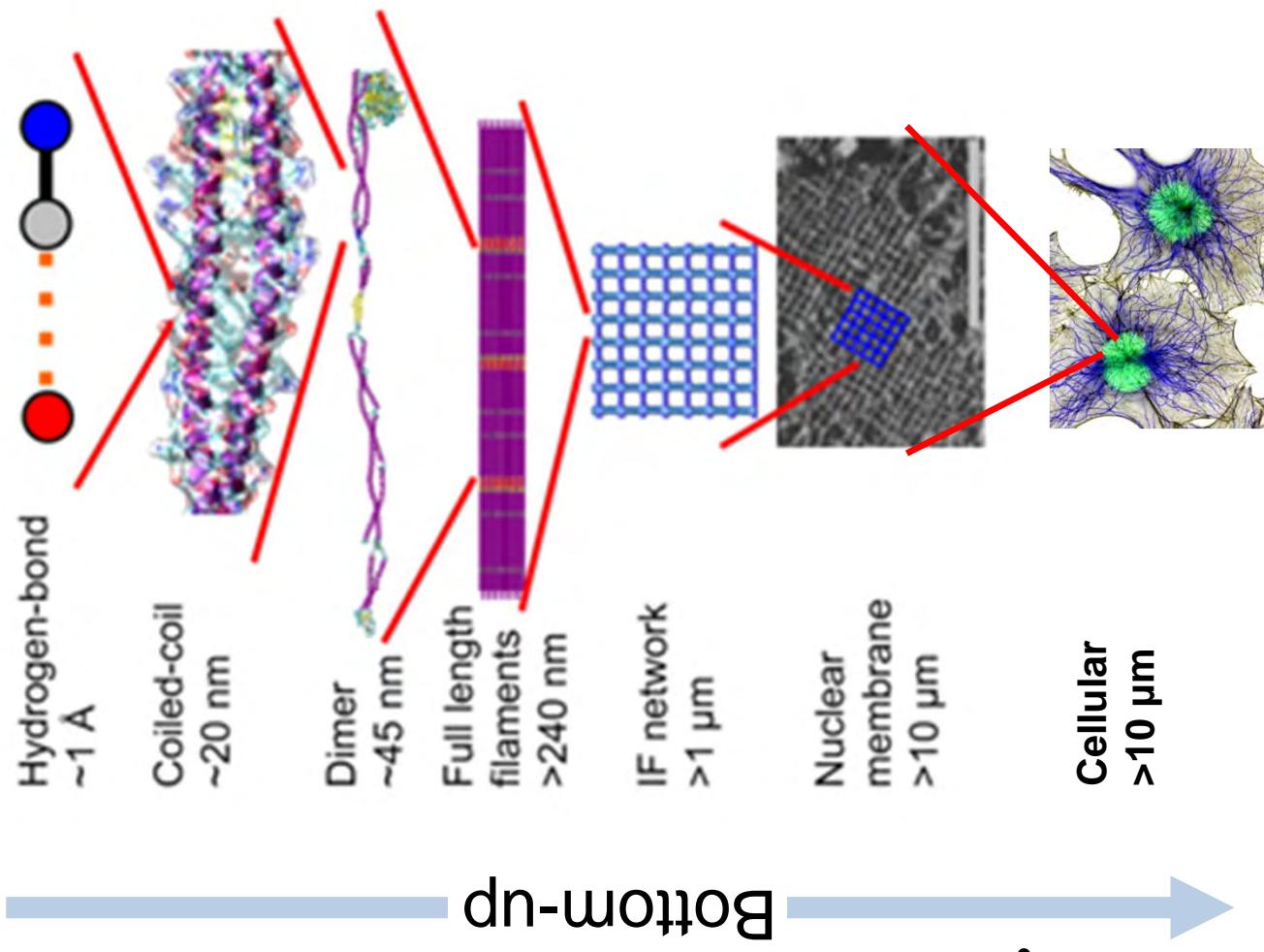


# Goals and benefits

- Identified **lamin intermediate filament protein material** in cells' nuclear lamina as model system to investigate key aspects of property-structural relation in biology
- Relevant properties: lamin meshwork is designed to protect the genetic material. Withstands **extreme mechanical deformation, strain stiffening at extreme deformation, strain stiffening at extreme deformation, very robustly** at varying conditions

## Goals and benefits

- Using the **bottom-up multi-scale model** of intermediate filaments to investigate the mechanism of material failure.
- Investigate the variant material property with **mutations** at the bottom level
- Design principles can be used in **design of de novo hierarchical nanomaterials**
- Expand engineering design space by merger of **structure property and function**.



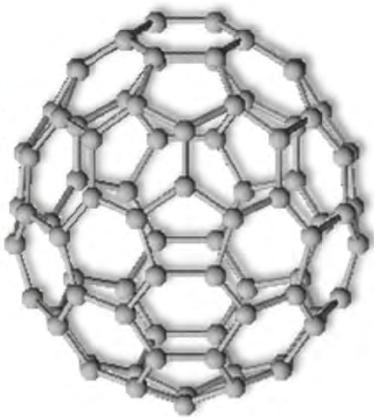
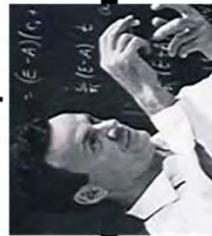


# Multiscale engineering

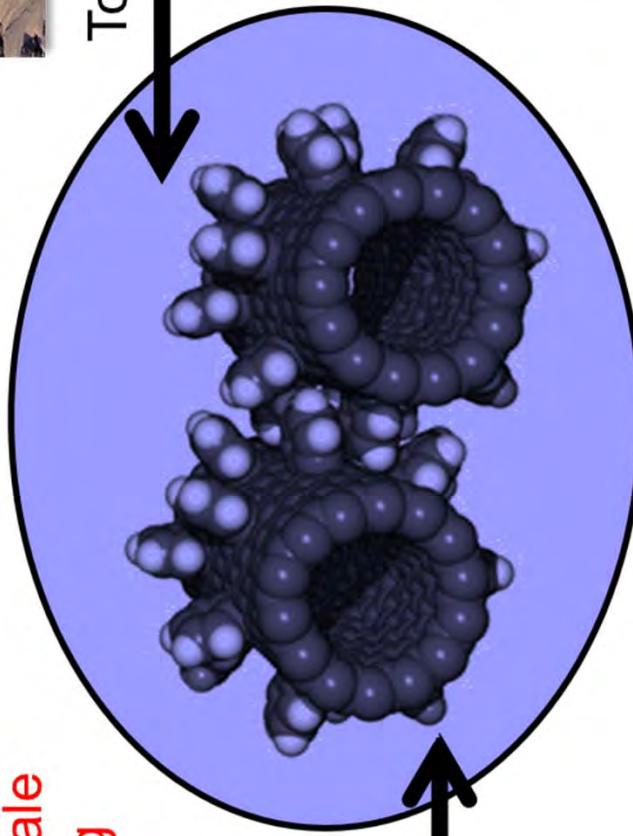
How much of the vast design space across scales is actually used by engineers? Not much!

**Opportunity: Multiscale science & engineering**

Nanoscience 1960s and following (Feynman paradigm)



Top-down (<BC  $\rightarrow$  2000s)



Closing the gap: now—**convergence of biology and engineering**

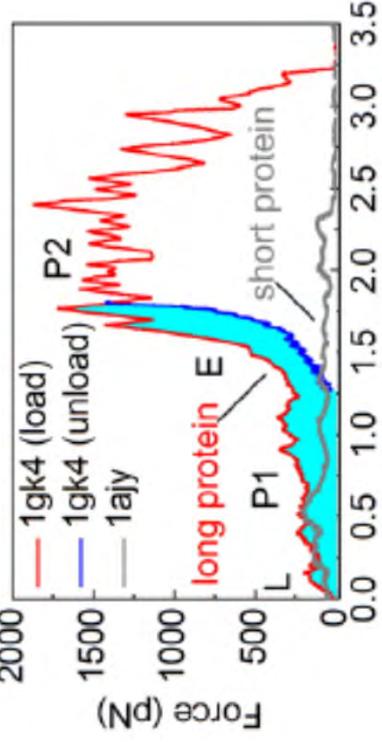
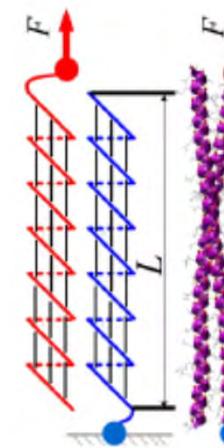
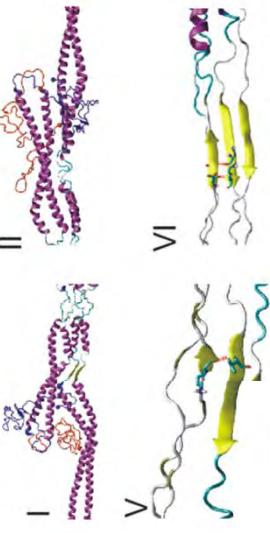


# Accomplishments

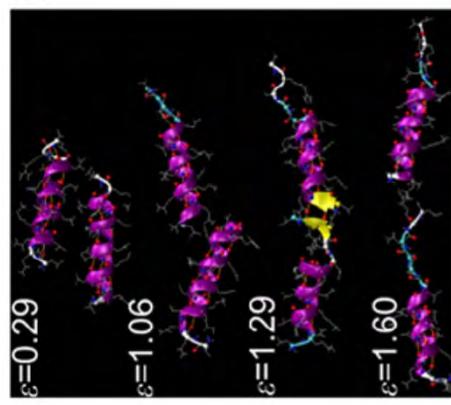
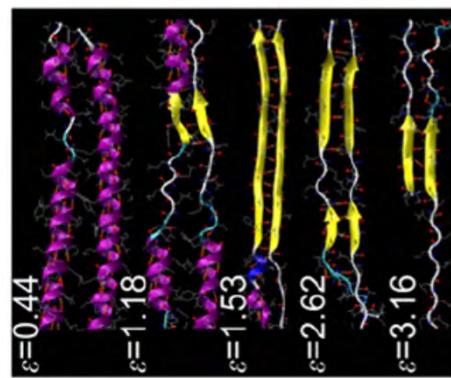
Predict and model alpha-to-beta transition as a key mechanism to control the stiffening character of helical protein materials, with important implications for larger-scale flaw-tolerant material properties.

- Alpha-to-beta transition mechanism is found for many coiled protein materials; emergence controlled by a characteristic length-scale (*not previously known*)

- Characterized nanomechanical and nanoscale behavior (modulus, strength, mechanism, filament thickness) based on molecular dynamics, validated against AFM experiment



The force-extension curve of a long coiled coil and a short coiled coil. The longer one features a stretching curve of beta sheet after transition

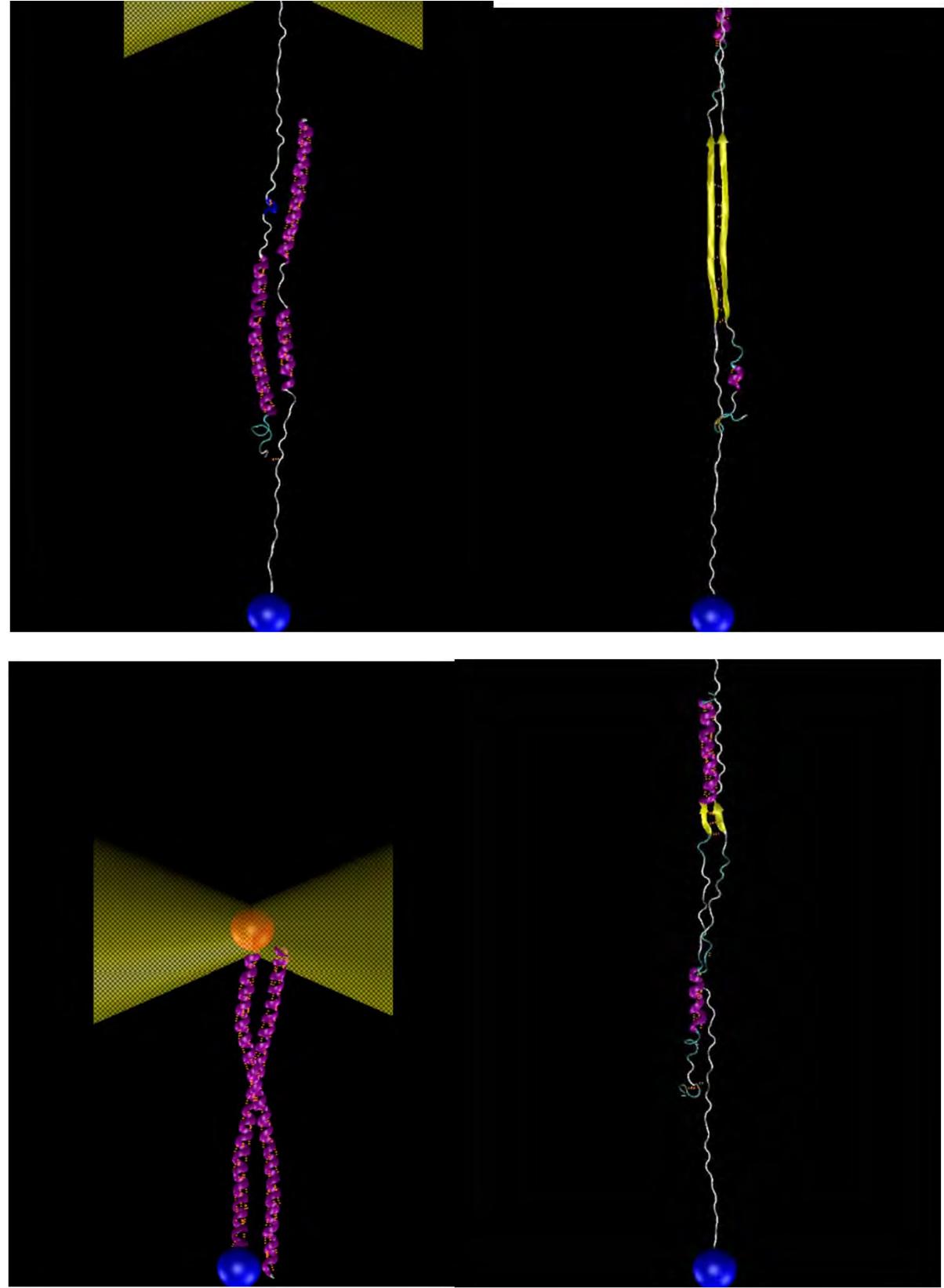
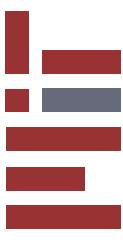


Short coiled coil without transition and long coiled coil with transition under deformation (full atomistic geometry)

Qin, Buehler, *PRL*, 2010



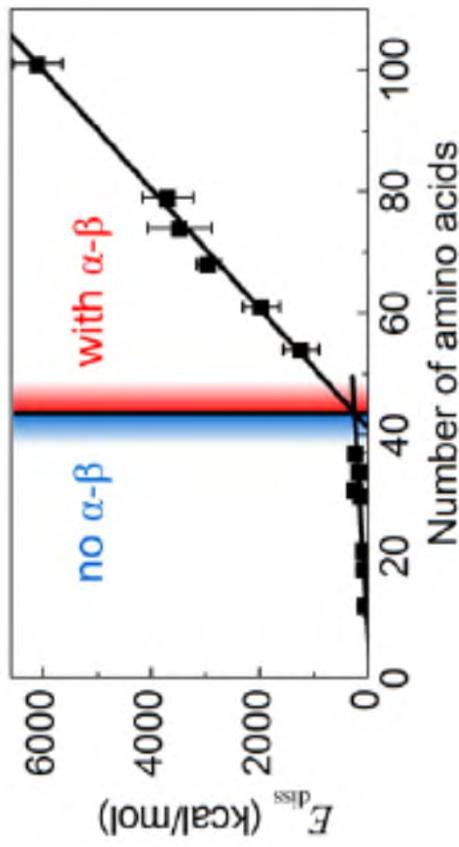
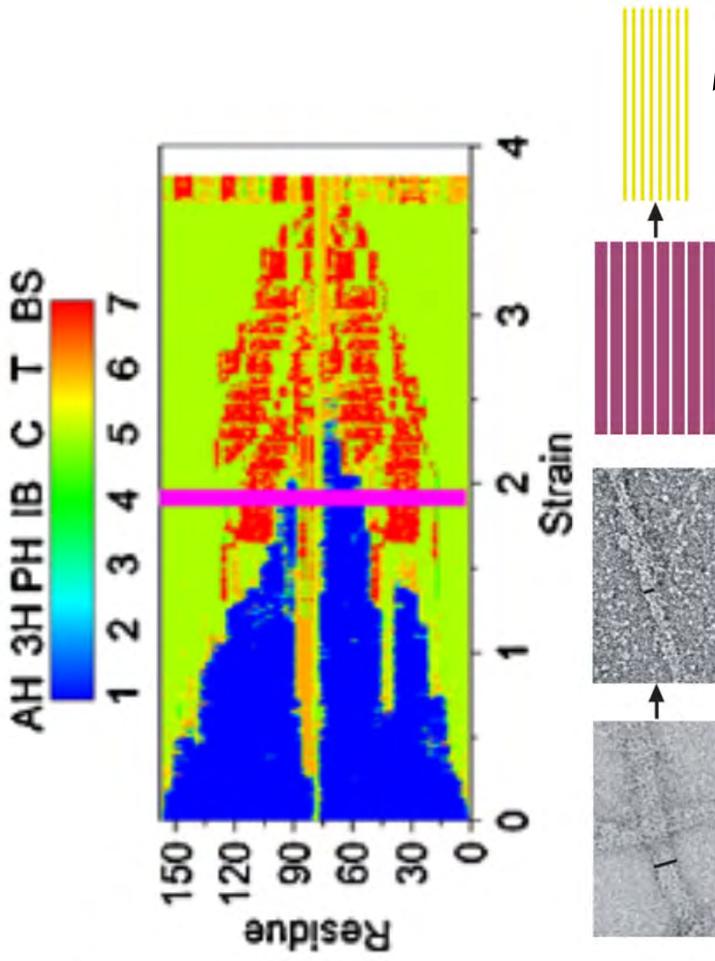
# Hierarchical mechanism in failure



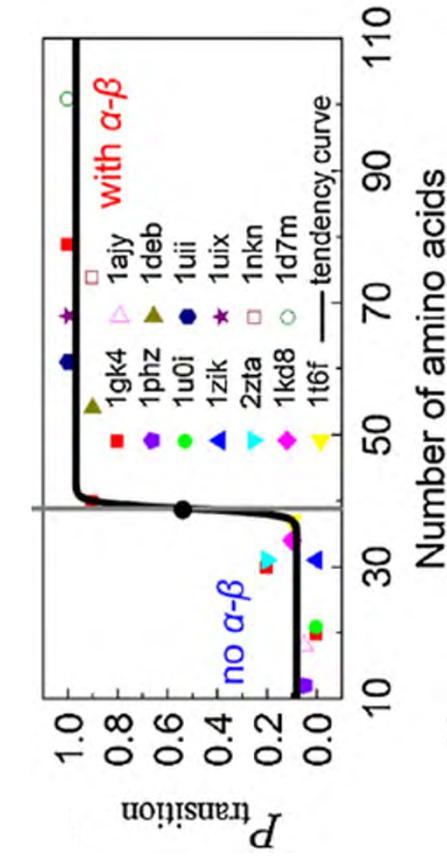


# Accomplishments

Structural transition during stretching of a coiled-coil structure (alpha helix: Blue; beta sheet: Red)



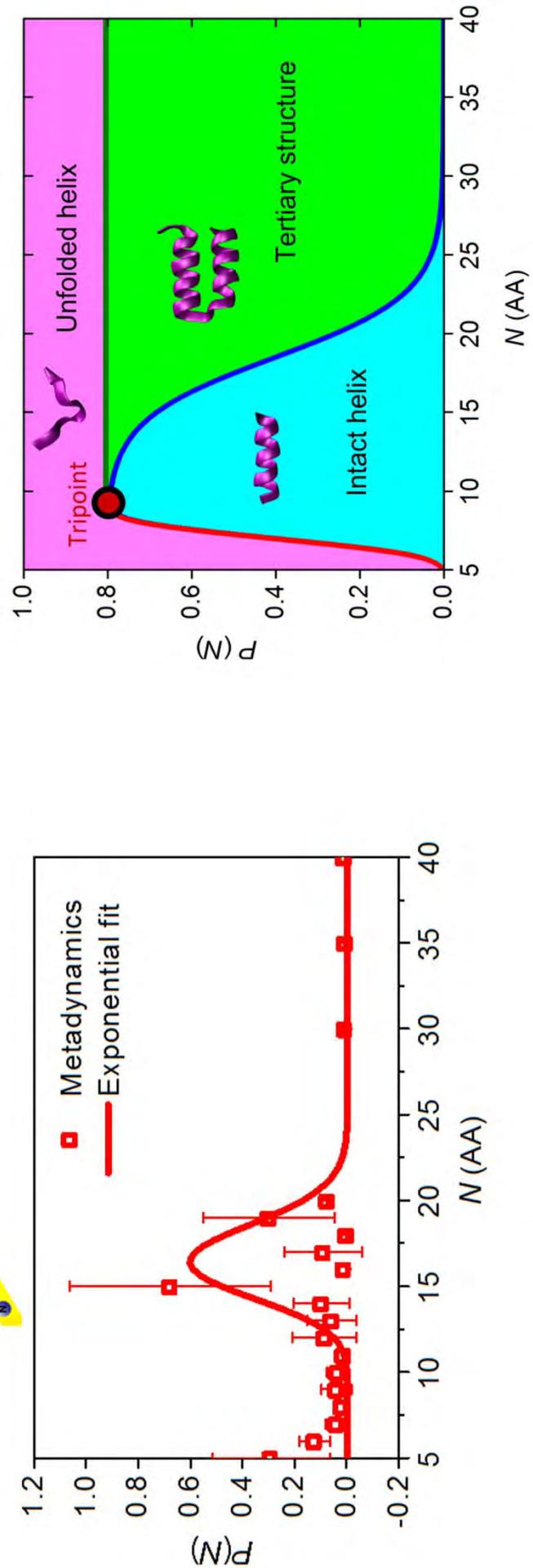
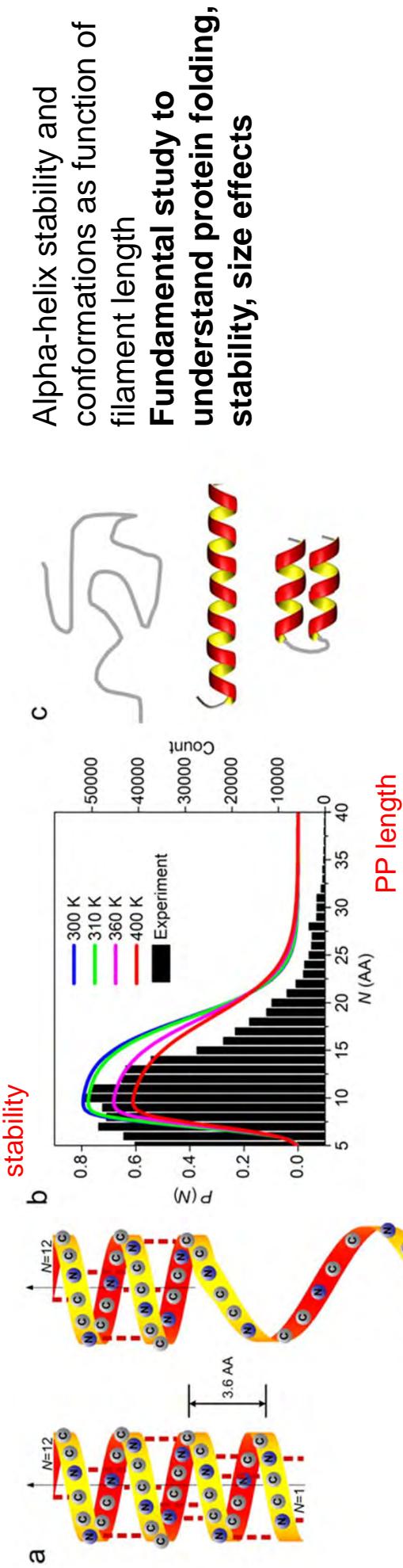
Validation against SEM experimental observation  
The alpha-to-beta transition increase the capacity of energy absorbtion



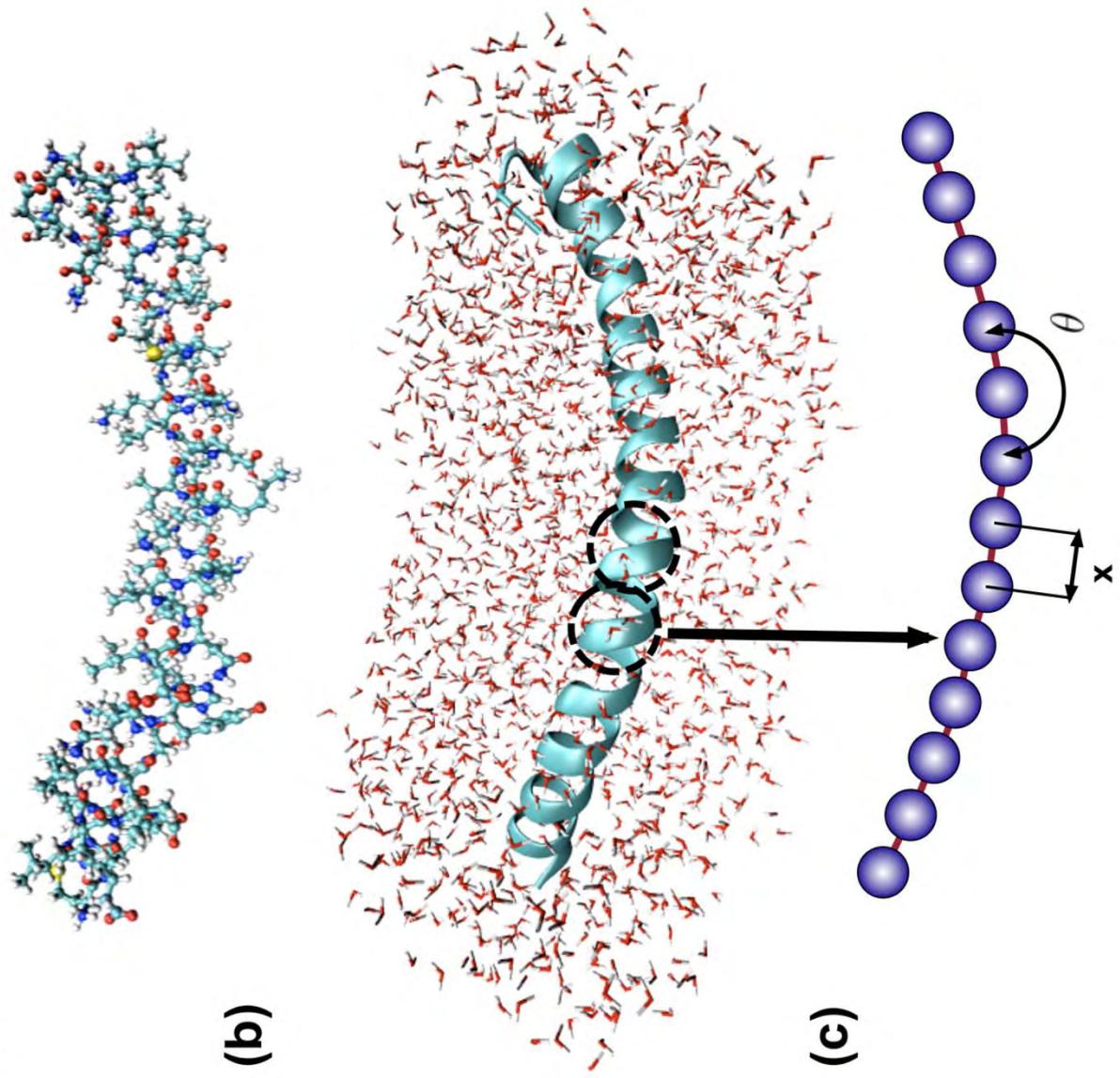
Critical length is found to govern the appearance of this transition process.



# Accomplishments

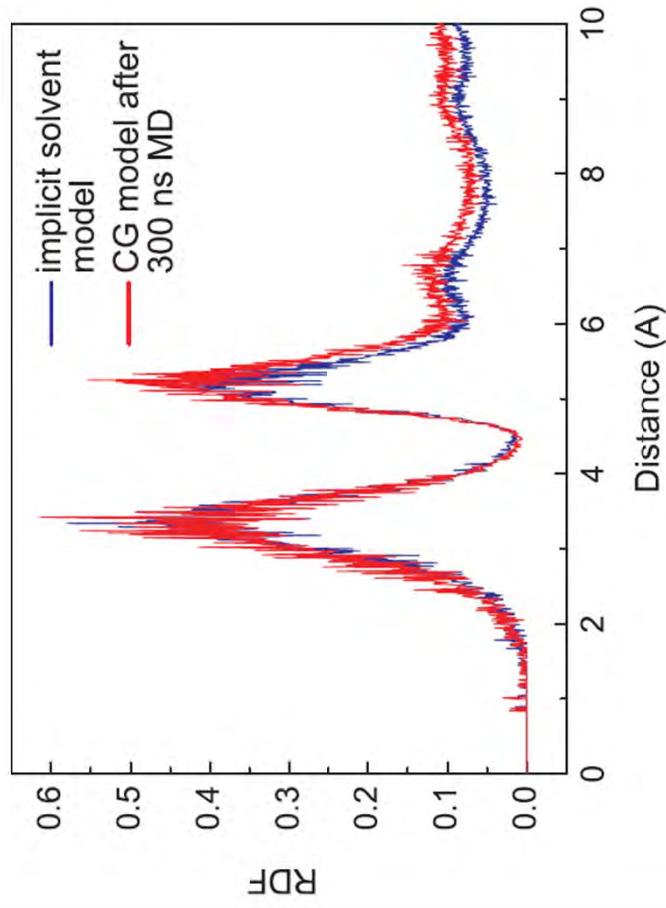
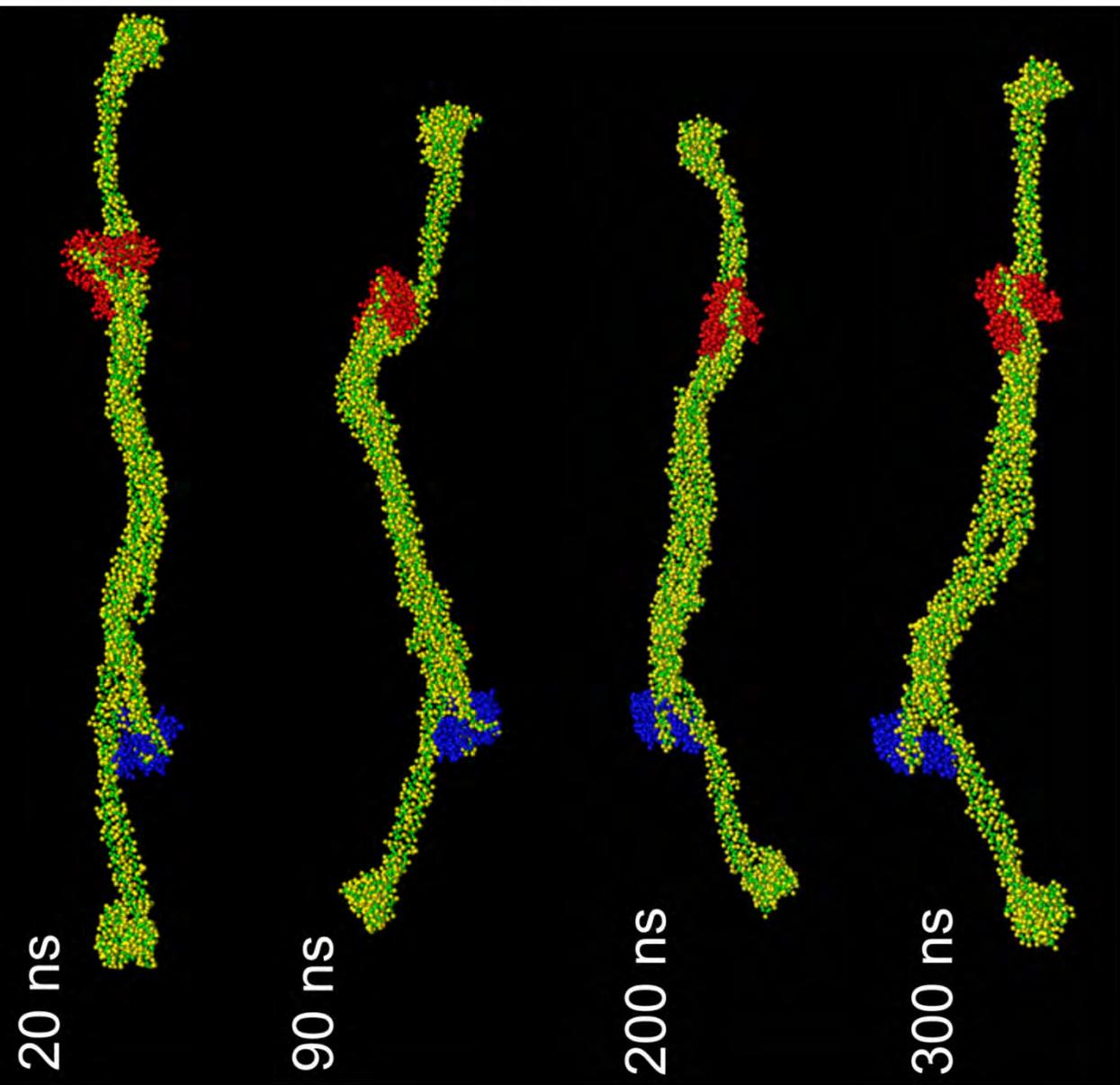


# Coarse-graining approach





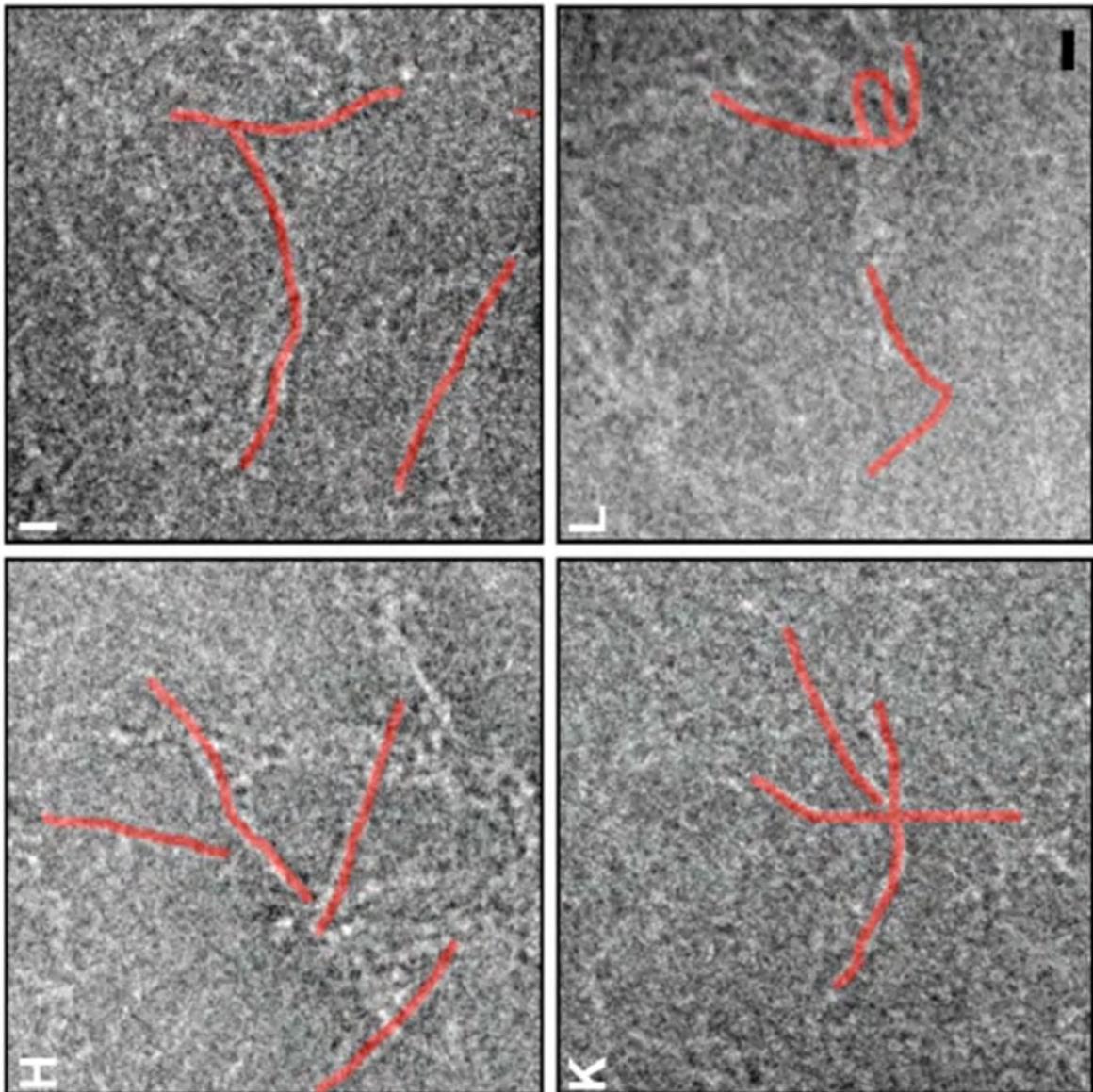
# Experimental validation





# Experimental validation (cryo-TEM)

cryo-negative stained desmin  
tetramers (experiment)

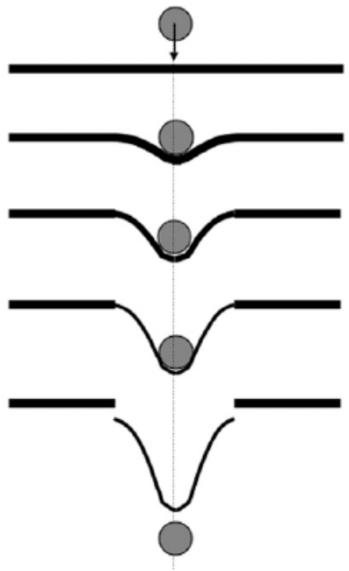
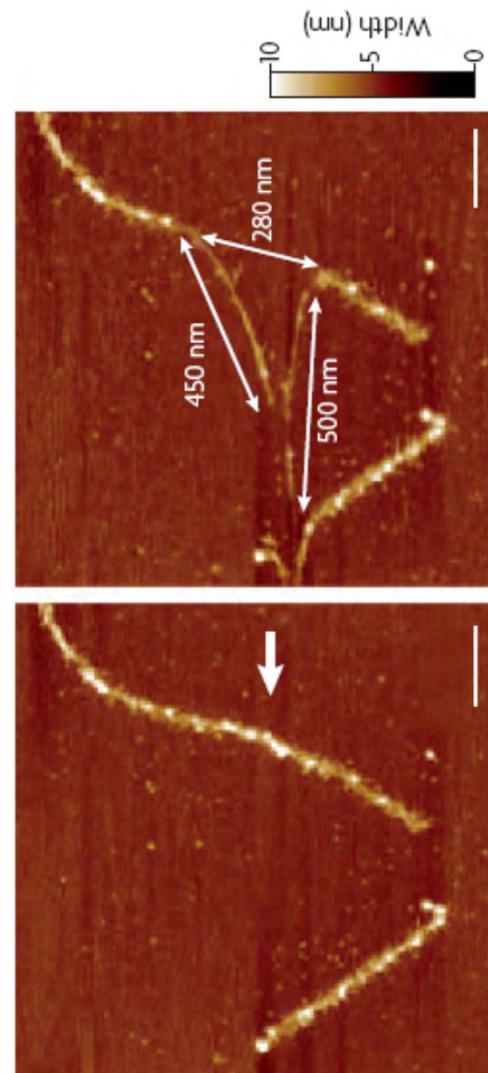
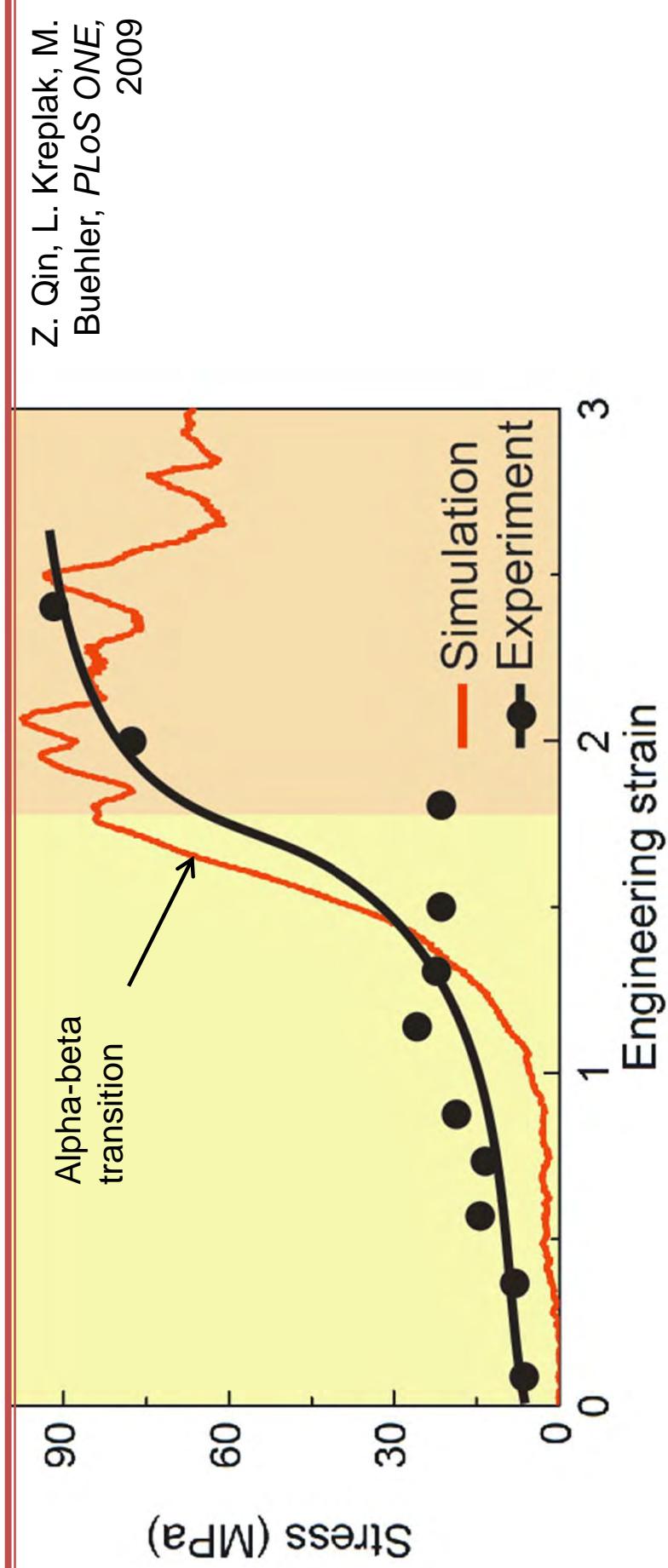


molecular  
model (theory)



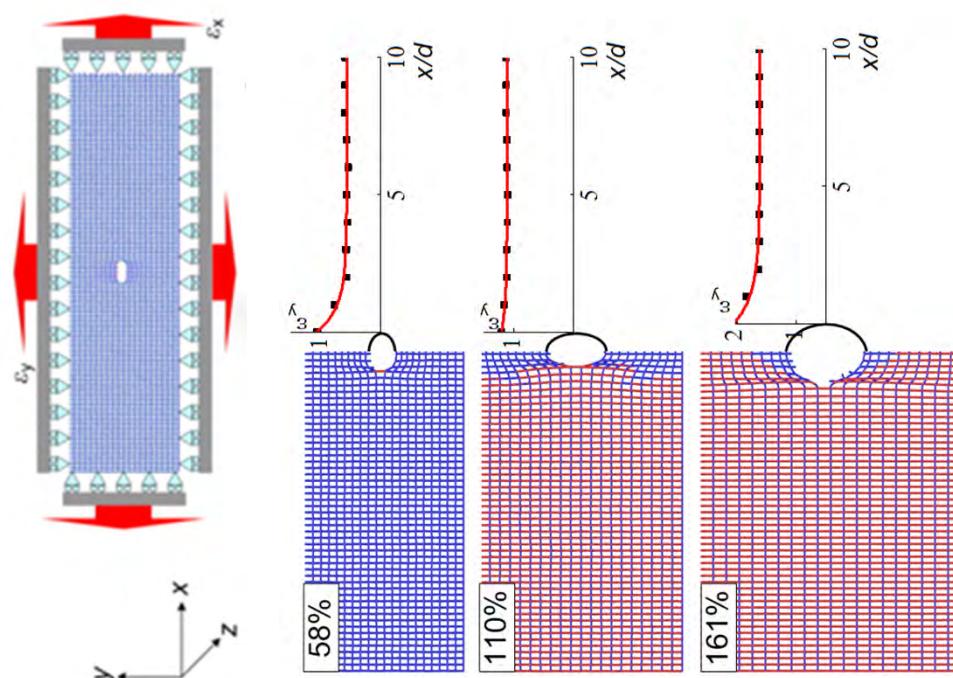
R. Kirmse, C.  
Weinert, Z. Qin,  
M. Buehler and  
L. Kreplak,  
*PLoS ONE*,  
2010

# Experimental validation (AFM)



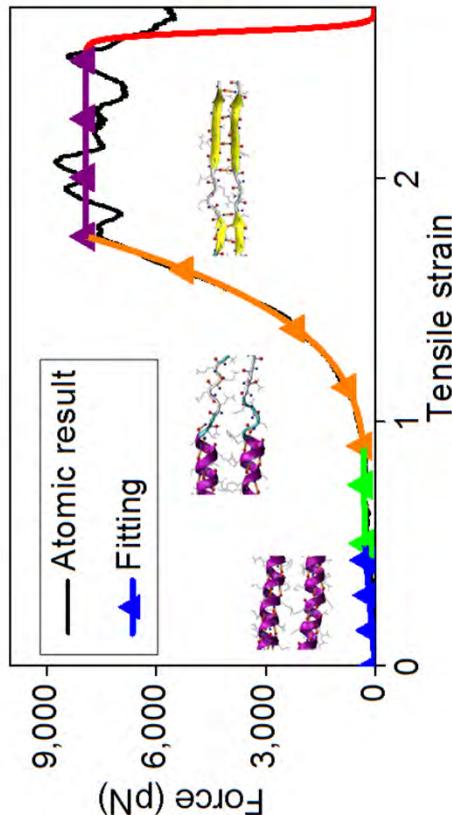


# Accomplishments



**Discover new *flaw-tolerance mechanism*:** Effect of non-linear property caused by *alpha-to-beta transition*.

Qin, Buehler, *in submission*



**Intact coiled coil** without defect, the deformation is elastic on hydrogen bonds.

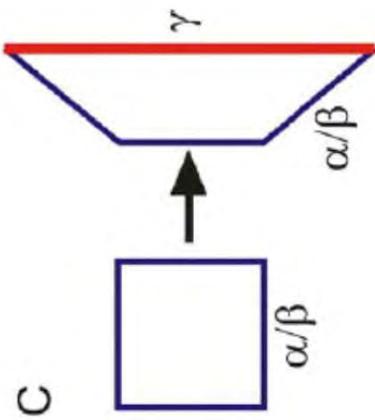
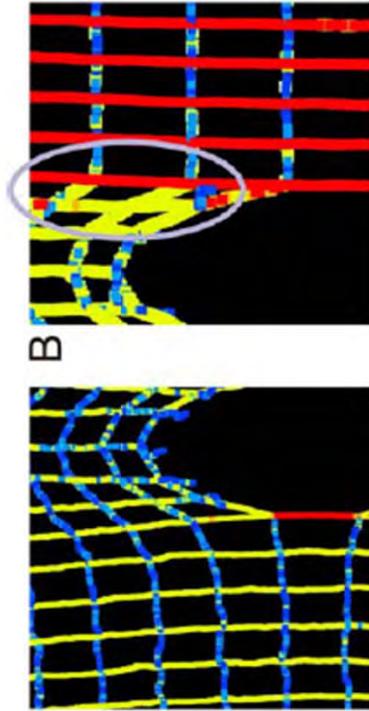
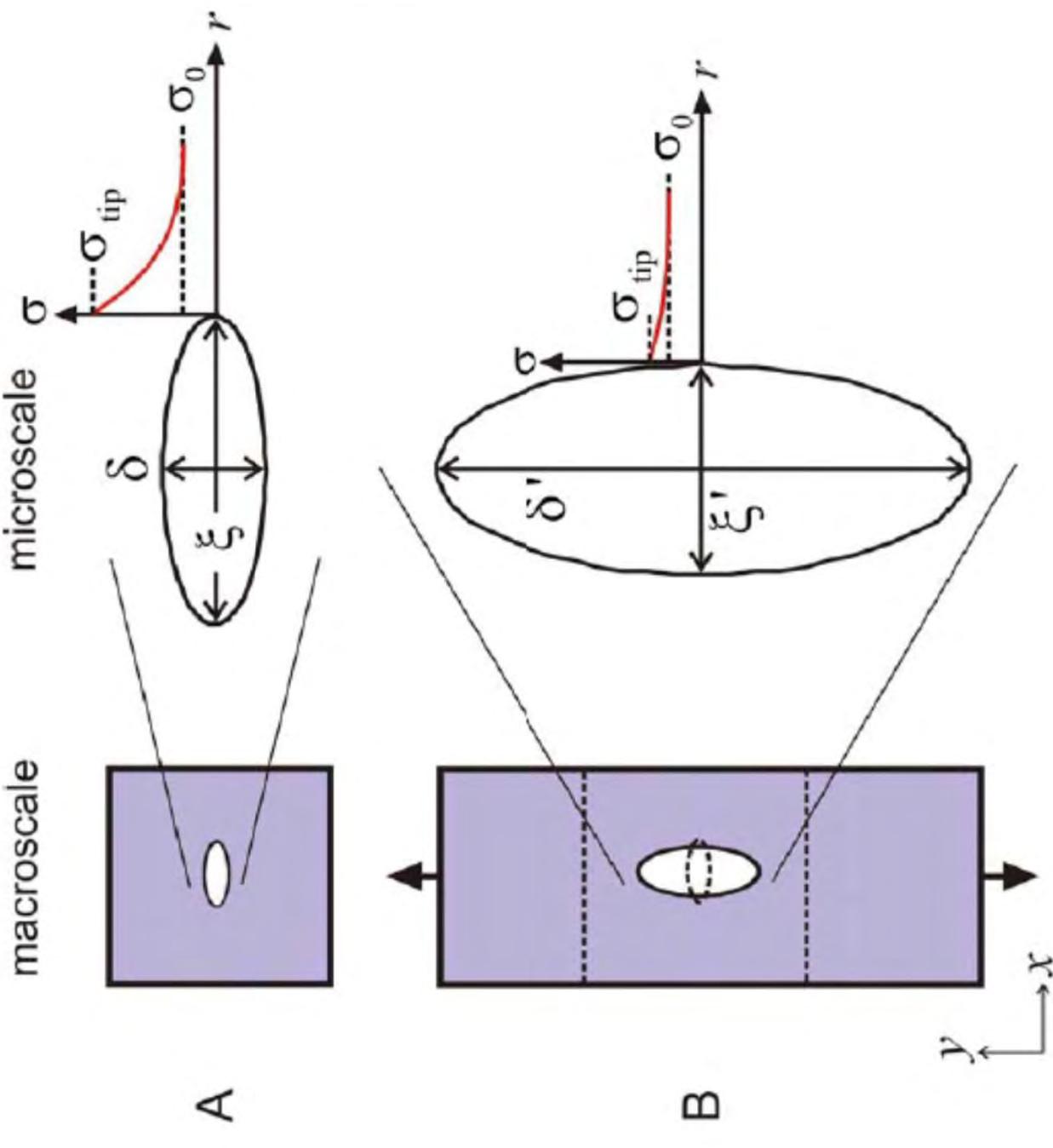
The coiled-coil structure **starts to unfold**, clusters of 3-4 hydrogen bonds rupture at the same time, providing a constant resistance against deformation.

The unfolded structure features the **alpha-to-beta transition**, making the material stiffening.

The beta-sheet strands **slides against** each other

The entire structure **falls apart**

# Accomplishments





# Accomplishments

Wild type

AA 646

Mature wildtype tail domain 386-646 (664 for pre-mature)

386 rrlrls

391 psptsqrsrg rasshssqtq gggsvtkkrk lestestrsff sqhartsgrv  
441 aveevdeegk fvrlnkrksnt dqsgmwnwqik rqngddpllt yrfppkftlk  
491 agqrvtiwa gagatshsppt dlwvqaqntw qgnnsrltal instgeevam  
541 rklyrsvtvv eddededord lnnhhhgshc sssgdppeyn lrartvlg  
591 cggpadkasa ssgsasvsv tvtrssyavg gsgggsgfdn  
641 lvtvryllgn ssprtqspan csim

AA 607

AA 614

Δ50 tail domain 386-614

386 rrlrls

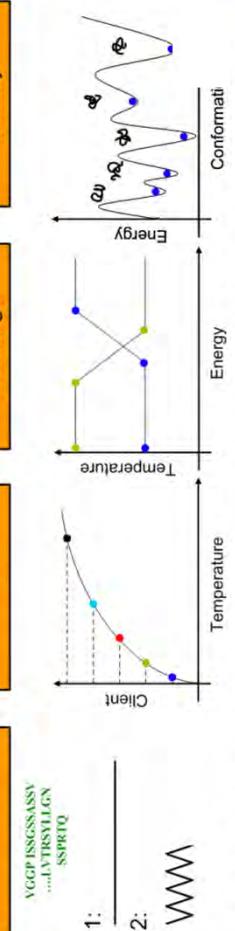
391 psptsqrsrg rasshssqtq gggsvtkkrk lestestrsff sqhartsgrv  
441 aveevdeegk fvrlnkrksnt dqsgmwnwqik rqngddpllt yrfppkftlk  
491 agqrvtiwa gagatshsppt dlwvqaqntw qgnnsrltal instgeevam  
541 rklyrsvtvv eddededord lnnhhhgshc sssgdppeyn lrartvlg  
591 cggpadkasa ssgsasvsv tvtrssyavg gsgggsgfdn  
641 lvtvryllgn ssprtqspan csim

• For nuclear lamina, study the structural and mechanical alteration after single point genetic mutation.

## Structure

- With protein folding (Replica exchange method), we manage to predict the structure of disordered region, which is not accessible for experimental crystallization (X-ray, NMR).

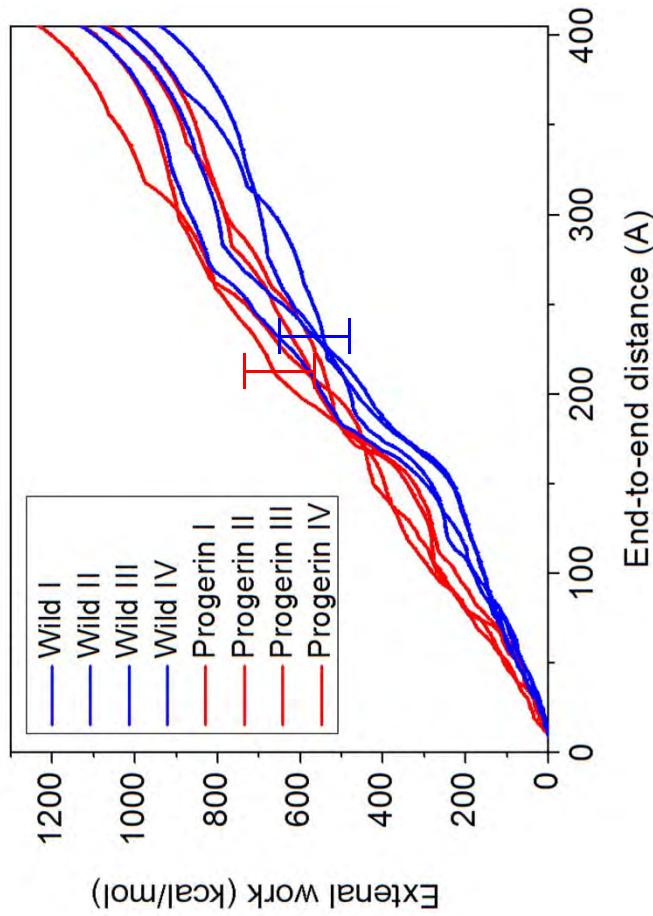
## Mechanics



- The point mutation in LAMN gene significantly alter the conformation of the tail domain of lamin, as well as the altered adhesion energy of this structure. The energy amounts agree with the measurement by experiments (tryptophan fluorometry, circular dichroism)



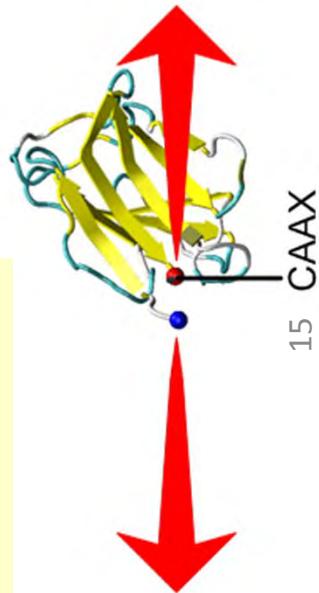
# Comparison with experiment



## In Simulation (SMD):

Extra energy is required to unfolding the Ig-fold in Progerin lamin

$$\Delta w = 69.8 \text{ kcal/mol}$$



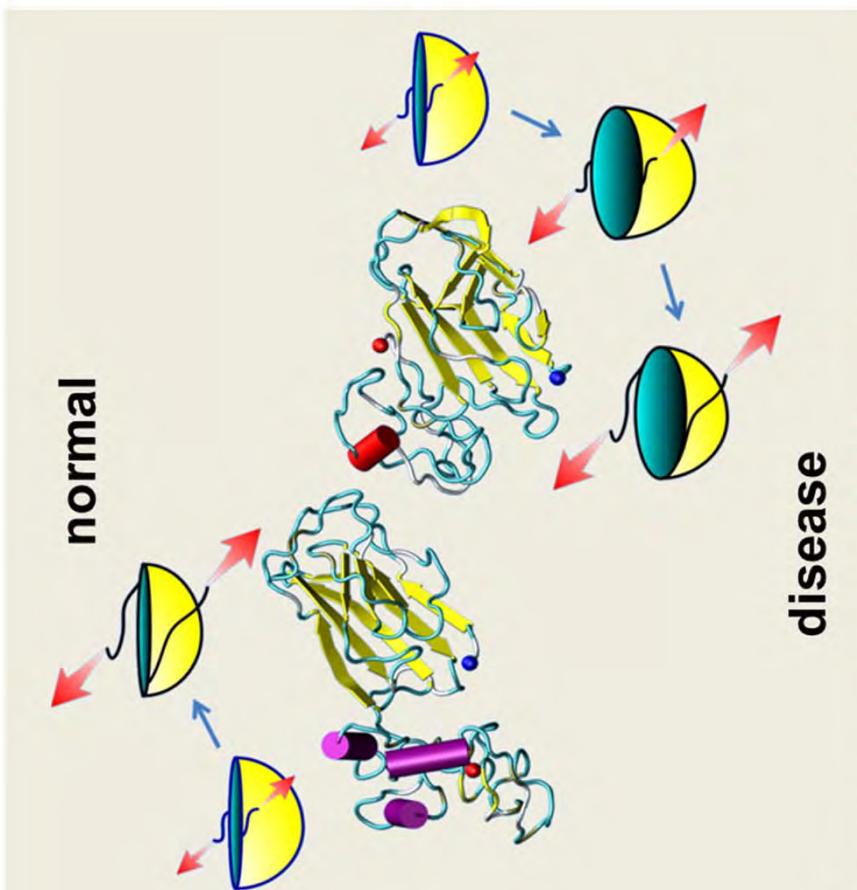
$$\text{Lam A } T_m = 61^\circ \text{ C} \uparrow$$
$$\text{progerin } T_m = 65^\circ \text{ C}$$

$$\Delta w = 56.2 \text{ kcal/mol}$$

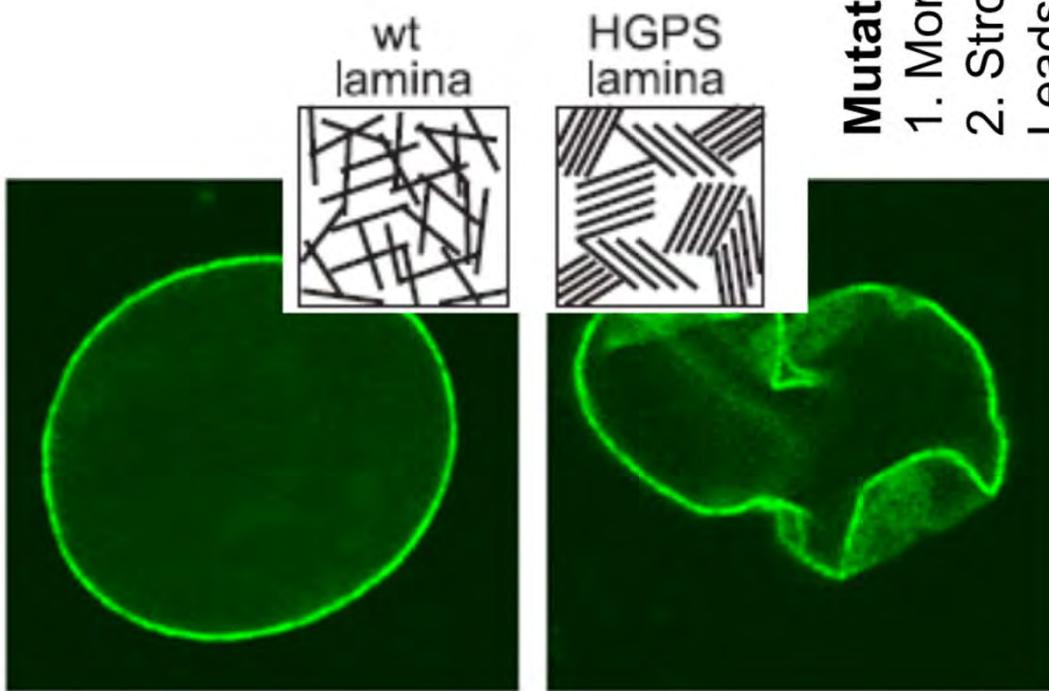
## In Experiment (Tryptophan fluorometry):

Higher temperature is required to unfolding the Ig-fold in Progerin lamin

# Comparison with experiment



Qin, Dahl, Buehler et al., *J. Struct. Biol.*, 2011



## Mutated structure:

1. More compact
2. Stronger interactions

Leads to “**embrittlement**” of nucleus; seen experimentally

**Multiscale model:**  
 $\Delta w = 69.8 \pm 17.4$   
kcal/mol

**Experiment:**

$\Delta w = 37 \pm 7.0$  kcal/mol

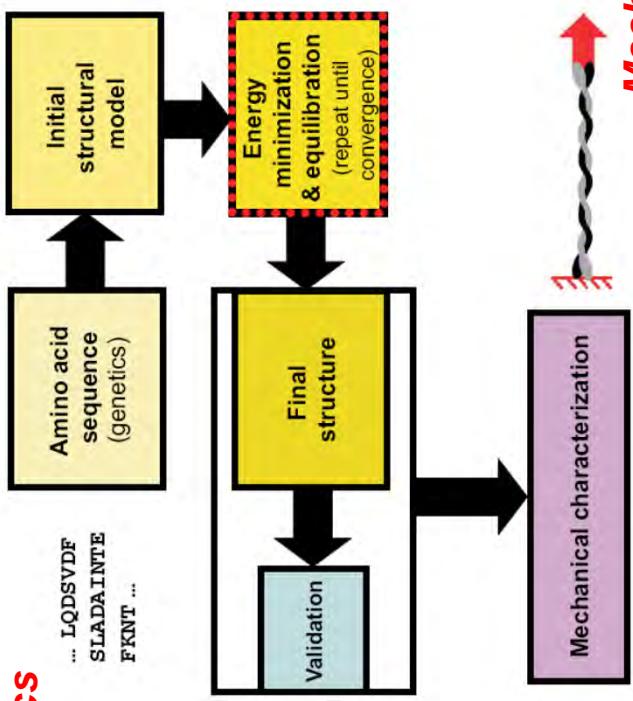


# Technical background

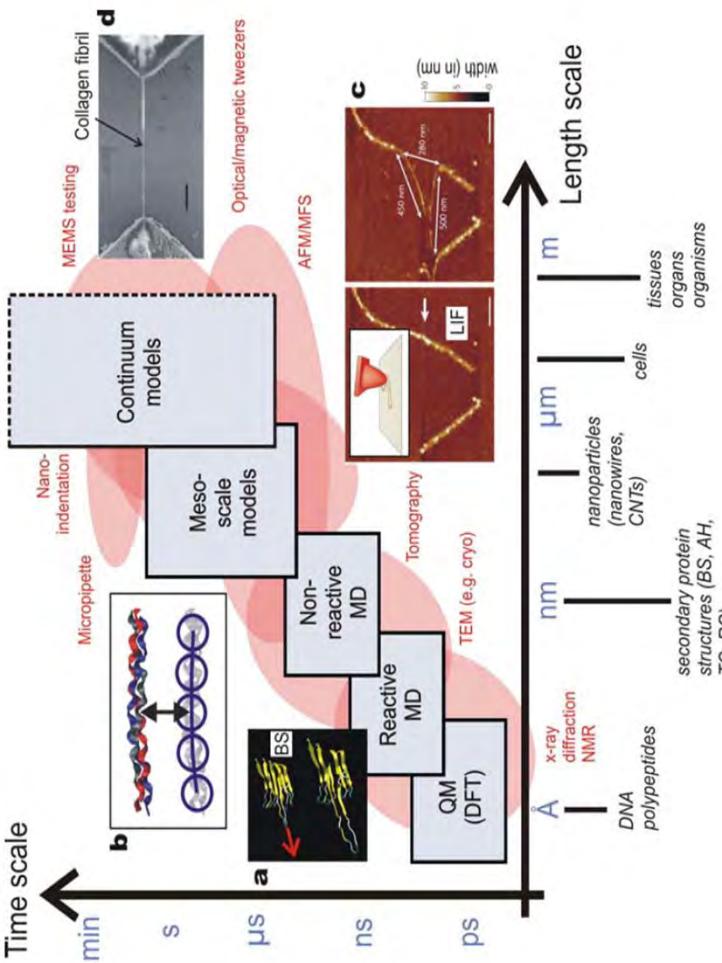
- **Molecular dynamics** based multi-scale simulation analysis of protein materials (atomistic and mesoscale modeling)
- Model assembly, deformation, failure
- **Theoretical models** (e.g. hierarchical Bell model) to develop structure-property relationships to provide design suggestions for biomimetic materials; **mechanistic insight**

## Genetics

... LQDSVDF  
SLADAINTE  
FKNT ...



**Materialomics:**  
Identify  
relationships  
between  
structure,  
properties and  
processes in  
hierarchical  
materials



# Research highlights

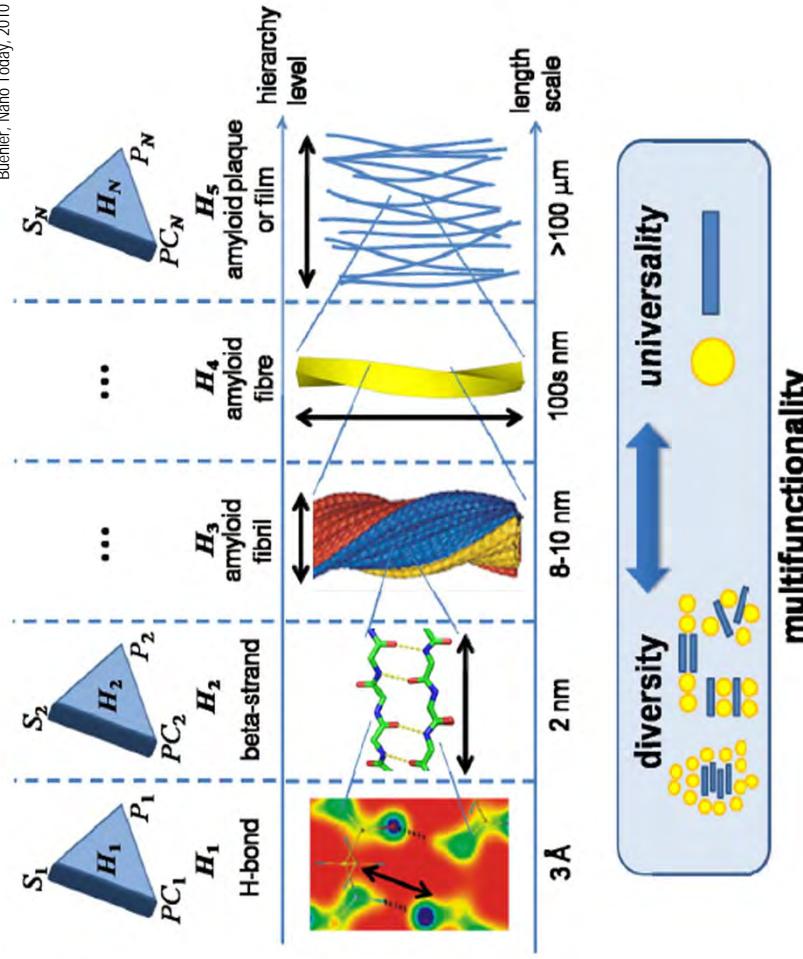


# Key points of research concepts

Understand **how biological materials are capable of combining disparate material properties** summarized below

Term	Definition
Strength	The maximum applied force or stress (pressure) at which failure of a system occurs (e.g. through fracture, tissue break down, etc.).
Robustness	Measures the ability of a system to tolerate flaws and defects, that is, still being capable of providing the required function under the presence of flaws and defects. A definition of robustness related to strength properties is the ratio of strength of a flawed structure divided by the strength of a perfect structure.
Adaptability	Ability of system to respond to changes in the environment (s.a. formation of defects due to injuries, or changes in physiological requirements, or due to the formation of fractures, etc.).
Flaw (defect)	Deviation of the structure of a system from its perfect, ideal or reference configuration. Examples for defects include cracks, inclusions, protein misfolds, or mutations in the amino acid sequence.
Failure	Sudden, typically uncontrolled and irreversible loss of the functional properties of a system. An example is the breakdown of tissue due to injuries under very large applied forces.
Self-healing ability	Ability of a system to reform from a perturbed structure to its reference configuration (reassemble). May involve for example the curing of flaws and defects such as cracks or voids, or the replacement or the addition of tissue.
Changeability and mutability	Formation of distinct (sometimes preprogrammed) structures with different properties, which can be controlled by external cues. Examples include the existence of multiple conformations of proteins based on pH, or applied forces.
Multifunctionality	Ability of a system to provide multiple properties to satisfy a set of target properties. An example is the combination of strength and robustness.
Evolvability	Ability of a system to evolve over generations of synthesis. In contrast to adaptability, evolvability reflects a change of structural makeup and/or properties over generations of synthesis.

Buehler, *Nature Nanotechnology*, 2010;  
Buehler, *Nano Today*, 2010



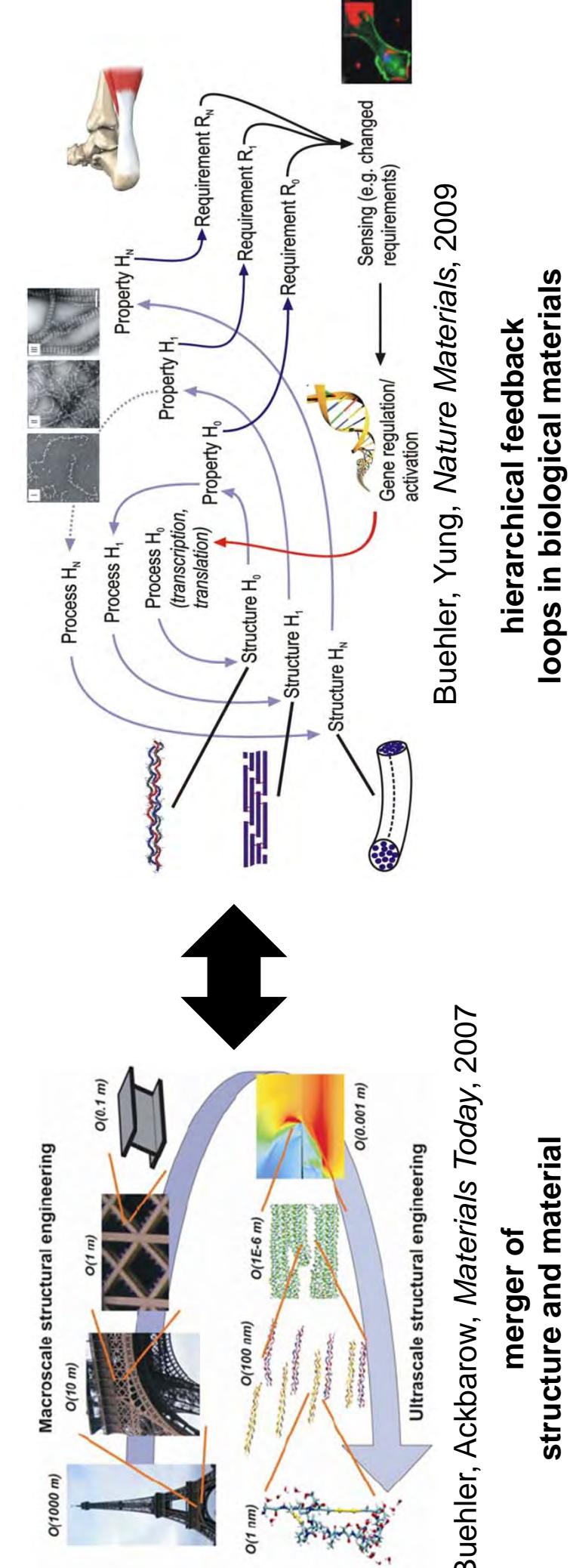
**Universality-diversity paradigm:** Create *multifunctionality (diversity)* by changing structural arrangements of few (*universal*) constituents rather than inventing new building blocks

Turn **weakness to strength**



# Expected benefits and future potentials

- Quantitatively understand the role of **hierarchies in materials** and their effect on material performance
- Through the analysis of biological structures, provide **suggestions for biomimetic and bioinspired material design**
- Provide **basis for new engineering paradigm** to **merge structure and material** (e.g. for new composites/heteropolymers with self-healing, mechanomutable capabilities): novel materials for advanced Air Force technologies



Buehler, Ackbarow, *Materials Today*, 2007

merger of  
structure and material

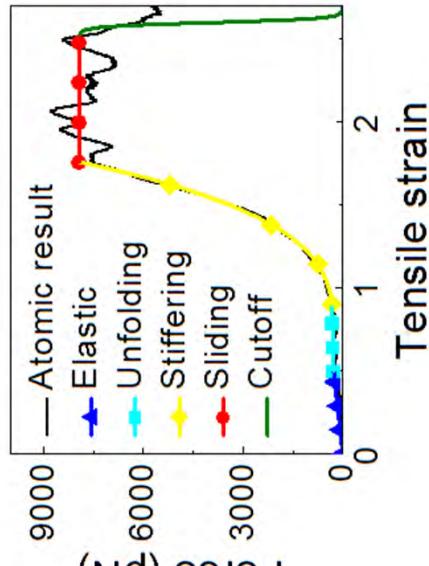
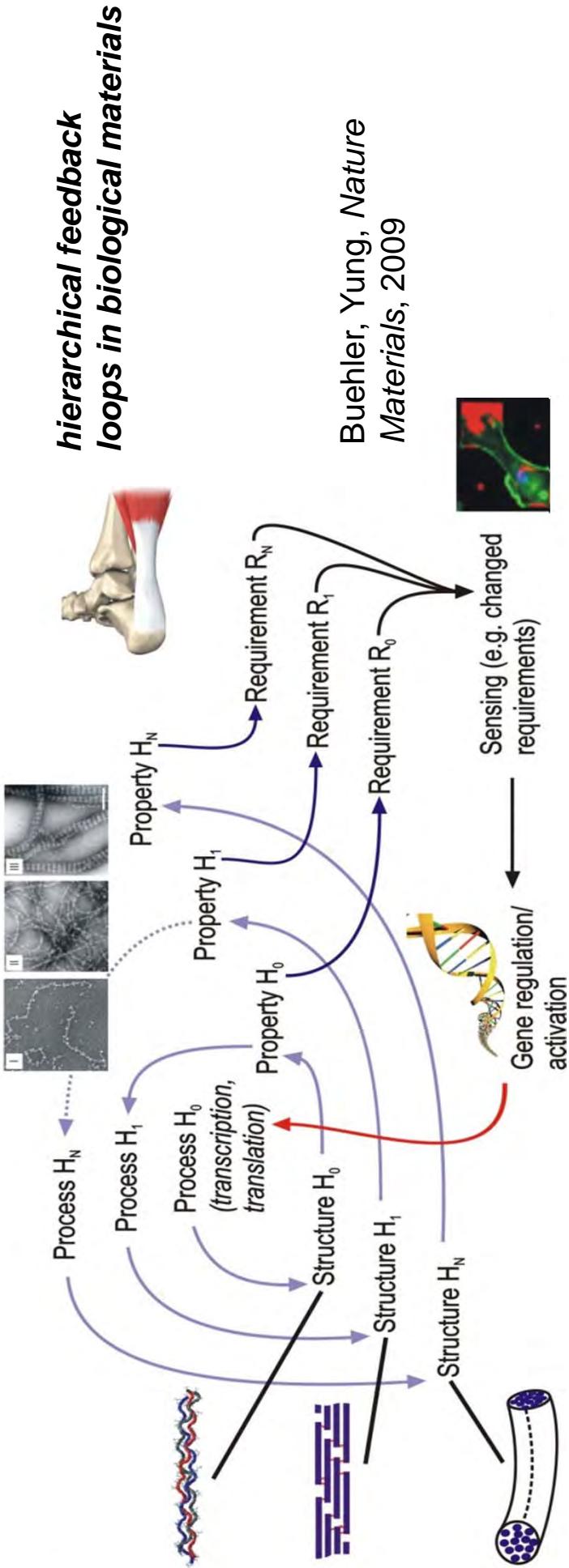
Buehler, Yung, *Nature Materials*, 2009

hierarchical feedback  
loops in biological materials



# Future study

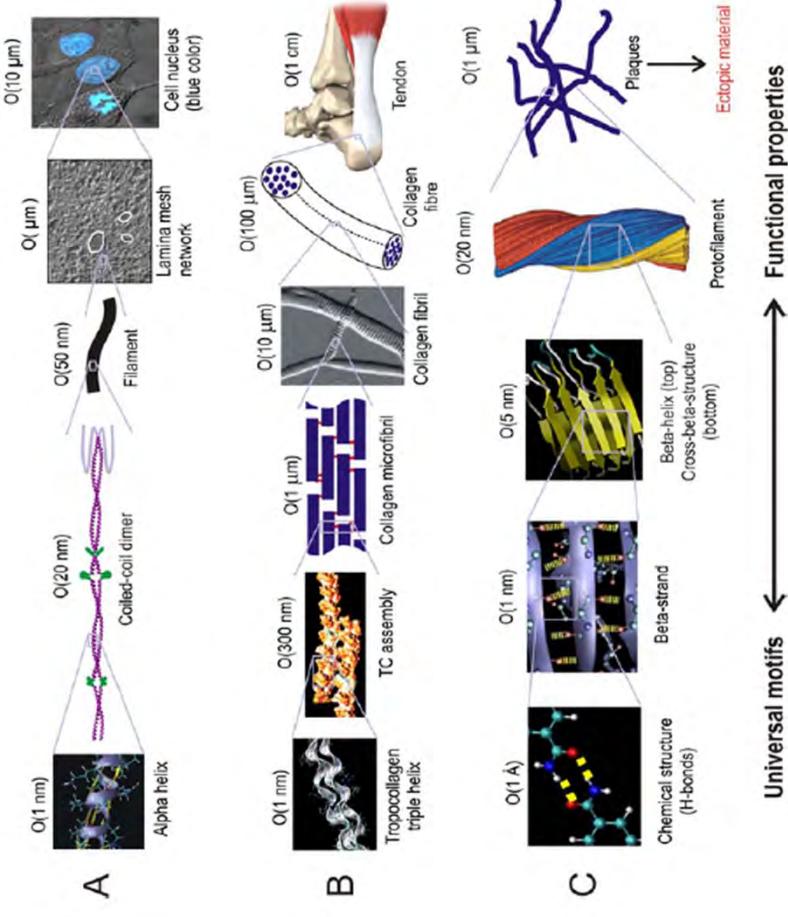
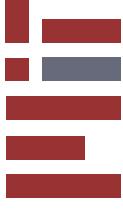
- Apply the mutation effect (with higher adhesive energy) to investigate the altered assembling process and mechanical property of full length filament.



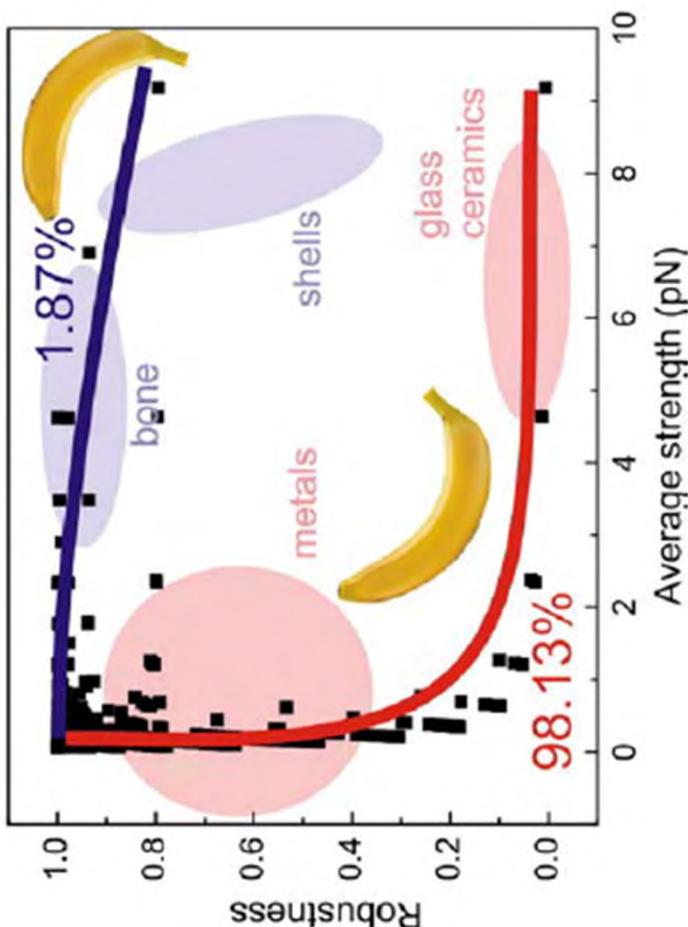
- Apply the altered mechanical and structural property to the meshwork level to investigate the effect of mutation on the modulus, strength and robustness of the protein material



# Technical background of research topic



Biological materials show hierarchical structures, as illustrated on the left for various materials on the left for various materials

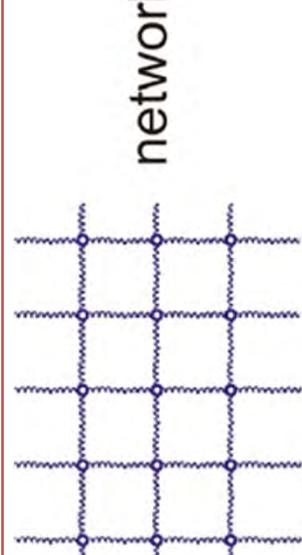
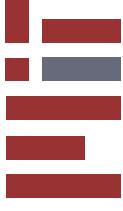


Biological materials are capable of combining disparate properties, such as strength and robustness, in a single material

Biological materials demonstrate how functional materials can be created out of simple and often inferior building blocks.



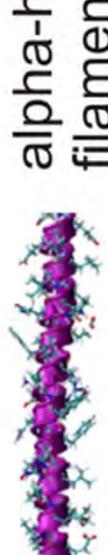
# Hierarchical mechanism in failure



**Structural transformation** of crack-like defects to mitigate stress concentrations



**Distance between filaments** facilitates extreme strain gradients at low energy cost



**Long linear array of turns** provides structural basis for large extensibility via repeated rupture of turns



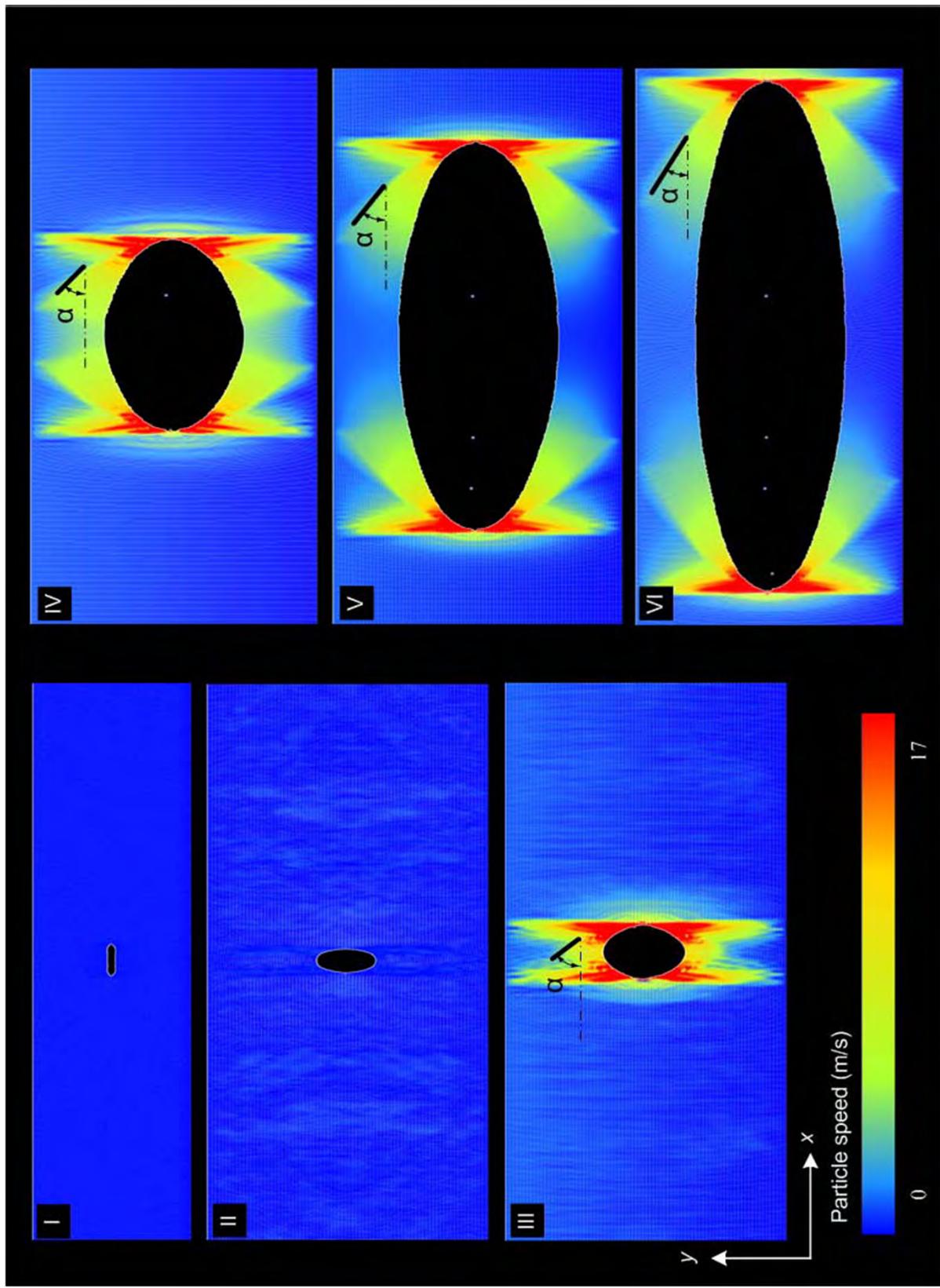
**Clusters of 3-4 H-bonds** provide optimal resistance against mechanical failure (3-4 H-bonds break concurrently)

HO H-bond

**H-bonds form at moderate temperatures: Self-assembly of alpha-helices/self-healing**

Ackbarow *et al.*, *PNAS*, 2007; Ackbarow, Sen, Thaulow, Buehler, *PLoS ONE*, 2009

# Hierarchical mechanism in failure





# Synthesis: Hierarchical function

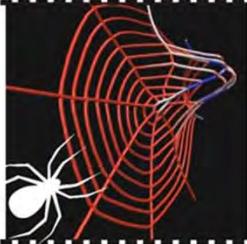


## protein



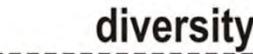
ACTG DNA/ACGU mR  
20 amino acids

spider web



hetero-nanocomposite

beta-sheet  
anocrystals



sensing & signalling  
feedback  
alter structure  
& property  
(mutability)

## universality

orchestral sound (symphony)

melody  
riff

-1

at different pitch intervals (compo-

modulating  
envelope  
(distinct)

100

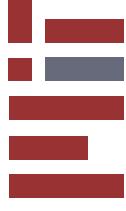
## music



M. Buehler, *Nano Today*, 2010



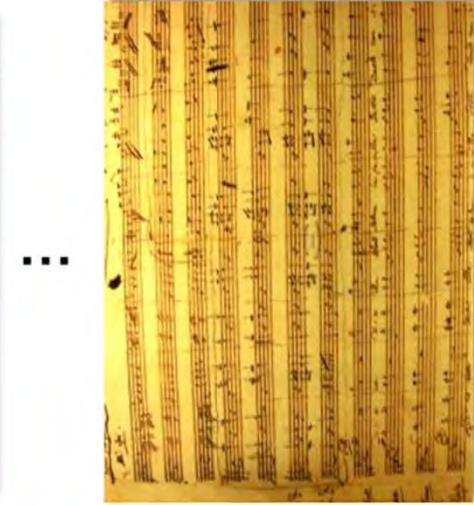
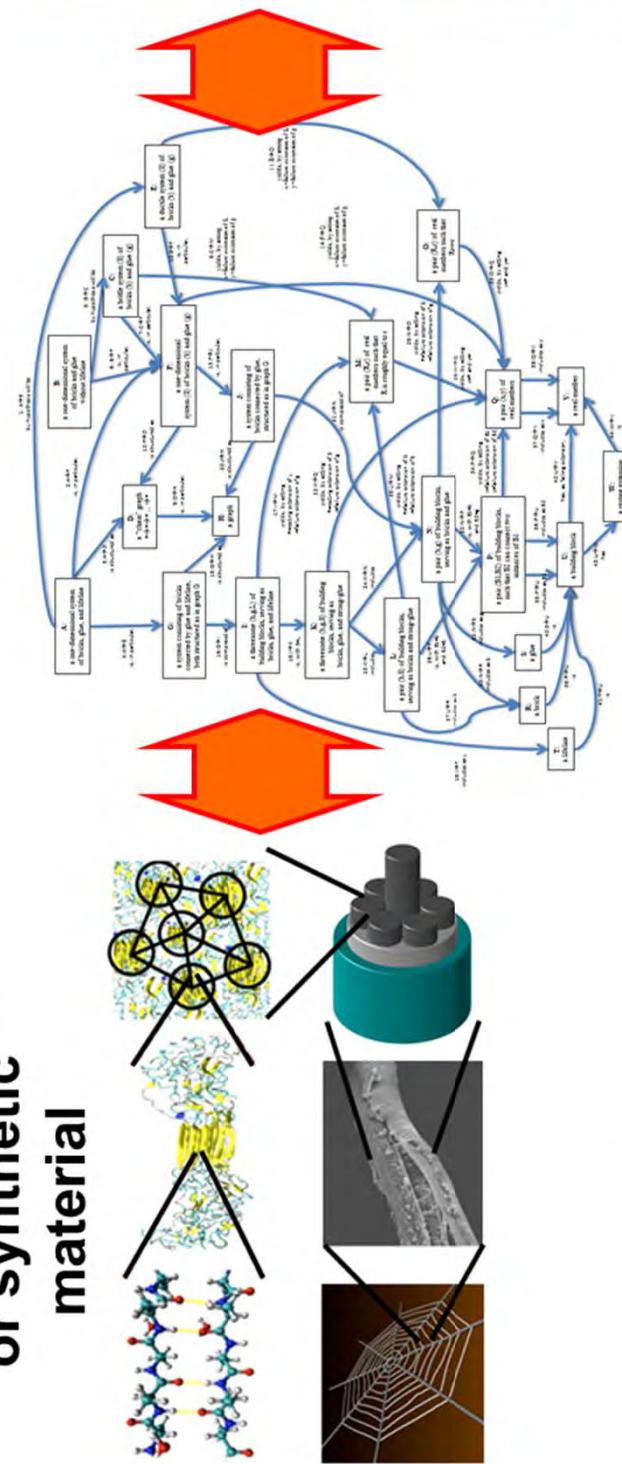
# Synthesis: Hierarchical function



## Social networks, music, language...

Description of how “function” emerges is identical/similar in silk or language, albeit building blocks are different

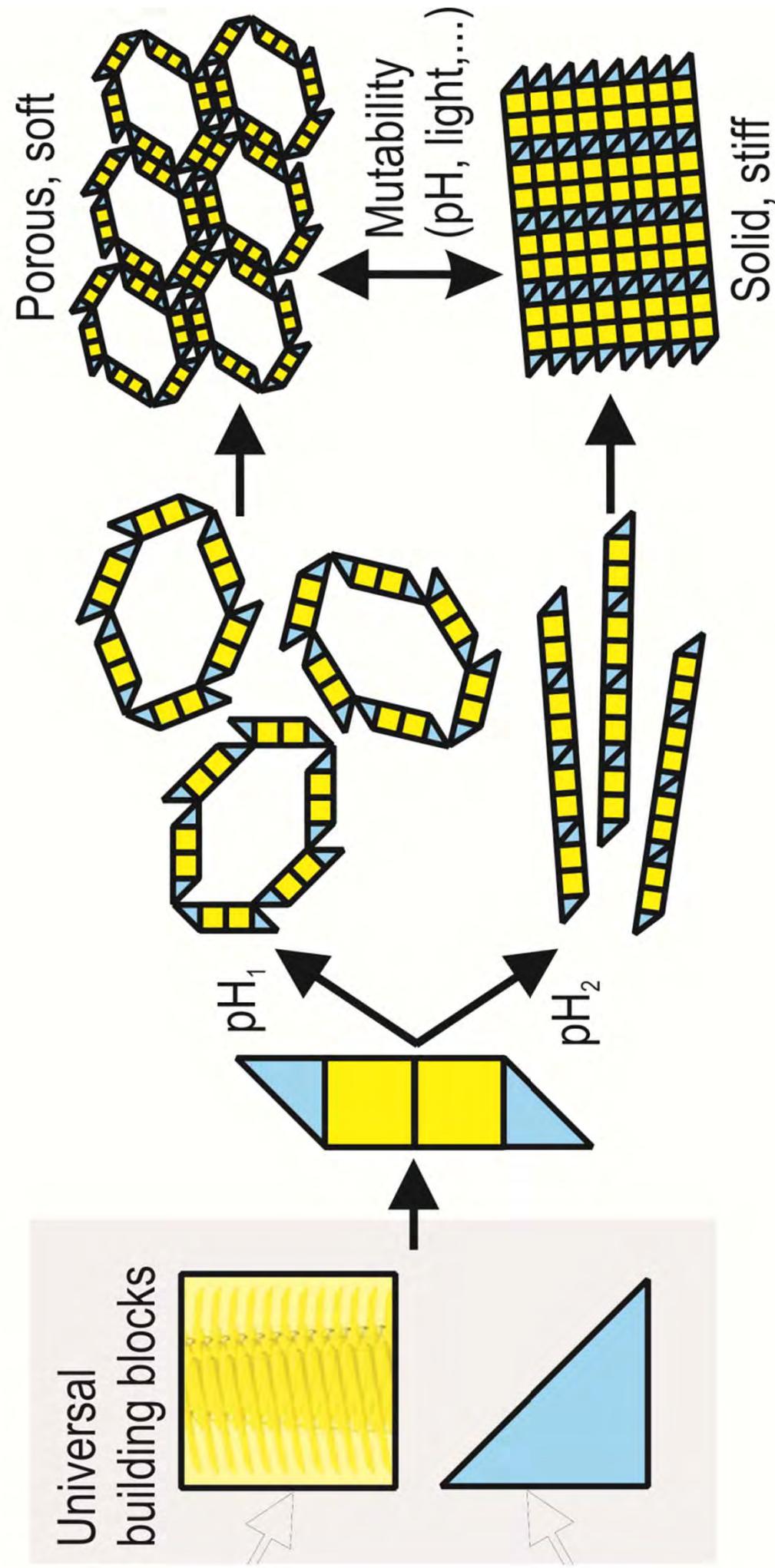
**Hierarchical  
biological  
material (e.g. silk)  
or synthetic  
material**



Solve olog for different “BCs”: change building blocks



# Hierarchical material design approach



# Summary of key accomplishments

- Developed first **atomistic-level molecular model** of human vimentin dimer and tetramer
- **Validated against experimental structure** analysis and bioinspired analysis/design
- Used molecular model to carry out **tensile testing experiments**, at dimer and tetramer level and upscaled to full length filament mechanics
- **Elucidated underlying mechanisms** (e.g. uncoiling of alpha-helices, alpha-beta-transition, interdimer sliding)
- **Future work:** **Improved structural model** more focus on **genetic diseases** – establish link **between genetics and biomechanics** (*muscle dystrophies, rapid aging disease progeria, etc...*)



# Awards received by the PI

- **Alfred Noble Prize** 2011 (presented by the combined engineering societies of the United States)
- Member, **Editorial Board UK Royal Society Interface**
- **Thomas J.R. Hughes Young Investigator Award** 2011 (for special achievements in Applied Mechanics for researchers under the age of 40, given by ASME's Applied Mechanics Division);
- **Chair, Fourth International Conference on Mechanics of Biomaterials & Tissues 2011**, Hawaii
- **Stephen Brunauer Award** 2011 (American Ceramic Society);
- **Leonardo Da Vinci Award** 2011 (Engineering Mechanics Institute of ASCE);
- **Rossiter W. Raymond Memorial Award** 2011 (AIME);
- **Best Paper Award**, International Journal of Applied Mechanics (Imperial College Press) 2010;
- **Sia Nemat Nasser Award** 2010 (given by ASME for research excellence in the areas of experimental, computational, and theoretical mechanics and materials);
- **Harold E. Edgerton Faculty Achievement Award** 2010 (highest honor bestowed upon young MIT faculty for exceptional distinction in teaching and research scholarship);
- **Presidential Early Career Award for Scientists and Engineers (PECASE)** 2009 (highest honor bestowed by the United States government on outstanding scientists and engineers beginning their independent careers).

## Key archived publications (2008-2011)

- Z. Qin, A. Fabre, M.J. Buehler, "Alpha-helical protein domains are most stable at a critical length scale," *Physical Review Letters*, in revision
- T. Knowles, M.J. Buehler, "Nanomechanics of functional and pathological amyloid materials," *Nature Nanotechnology* 6(7), 469–479, 2011
- Z. Qin, A. Kalinowski, K. Dahl, M.J. Buehler, "Structure and stability of the lamin A tail domain and HGPS mutant," *Journal of Structural Biology* 175, 425–433, 2011.
- Z. Qin, M.J. Buehler, "Dynamic failure of nuclear lamina under extreme mechanical deformation," *BioNanoScience*, 1(1), 14–23, 2011
- M. Arslan, Z. Qin, M. Buehler, "Coiled-coil intermediate filament stutter instability and molecular unfolding," *Computer Methods in Biomechanics and Biomedical Engineering*, 14(5), 483–489, 2011
- Z. Qin, M.J. Buehler, "Flaw tolerance of nuclear intermediate filament lamina under extreme mechanical deformation," *ACS Nano* 5(4), 3034–3042, 2011
- Zhao Qin, Markus J. Buehler (2011). Structure and dynamics of human vimentin intermediate filament dimer and tetramer: Comparison between explicit and implicit solvent molecular dynamics, *J. Mol. Model.* 16(10) : 2940
- Z. Qin, M.J. Buehler, "Cooperative deformation of hydrogen bonds in beta-strands and beta-sheet nanocrystals," *Physical Review E*, 82(6), paper # 061906, 2010
- R. Kirmse, Z. Qin, C.M. Weinert, A. Hoenger, M.J. Buehler, L. Kreplak, "Plasticity of intermediate filament subunits," *PLoS ONE* 5(8), paper # e12115, 2010
- Markus J. Buehler, Zhiping Xu (2010). Mind the helical crack, *Nature*, 464(4):42–43.
- Zhao Qin, Markus J. Buehler (2010). Critical length scale governs alpha-beta transition in coiled-coil protein filaments, *Physical Review Letters*, 104: 198304
- Sinan Keten, Zhiping Xu, Britni Ihle, Markus J. Buehler (2010), Nanoconfinement controls stiffness, strength and mechanical toughness of beta-sheet crystals in silk, *Nature Materials*, 9:359–367
- Markus J. Buehler, Sinan Keten (2010). Failure of molecules, bones, and the earth itself, *Rev. Mod. Phys.*, 82(2).
- Markus J. Buehler (2010), Strength in numbers, *Nature Nanotechnology*, 5:172–174
- Markus J. Buehler, Yu Ching Yung (2010), How protein materials balance strength, robustness and adaptability, *HFSP Journal*, 4(1):26–40.
- Jérémie Bertaud, Joshua Hester, Daniel D. Jimenez, Markus J. Buehler (2010). Energy landscape, structure and rate effects on strength properties of alpha-helical proteins, *J. Phys. Condensed Matter*, 22, paper # 035102.
- Jérémie Bertaud, Zhao Qin and Markus J. Buehler (2010). Intermediate filament deficient cells show mechanical softening at large deformation, *Acta Biomaterialia* 6 : 2457-2466
- Zhao Qin, Laurent Kreplak, Markus J. Buehler (2009). Hierarchical Structure Controls Nanomechanical Properties of Vimentin Intermediate Filaments. *PLoS ONE*, 4 : e7294
- Zhao Qin, Laurent Kreplak, Markus J. Buehler (2009). Nanomechanical properties of vimentin intermediate filament dimers, *Nanotechnology*, 20 : 425101
- Zhao Qin, Markus J. Buehler, Laurent Kreplak (2010). A multi-scale approach to understand the mechanobiology of intermediate filaments, *J. Biomechanics*, 43 : 15–22
- Jérémie Bertaud, Zhao Qin and Markus J. Buehler (2009). Atomistically Informed Mesoscale Model of Alpha-Helical Protein Domains, *I. J. Mult. Comp. Eng.*, 7: 237–250
- Jérémie Bertaud, Zhao Qin and Markus J. Buehler (2009). Amino acid sequence dependence of nanoscale deformation mechanisms in alpha-helical protein filaments. *J. Strain Analysis*, 44 : 517–531
- Theodor Ackbarow, Dipanjan Sen, Christian Thaulow, Markus J. Buehler (2009). Alpha-Helical Protein Networks are Self Protective and Flaw Tolerant, *PLoS ONE*, 4(6), paper # e6015.
- Ackbarow, S. Keten, and M.J. Buehler, "A multi-timescale strength model of alpha-helical protein domains", *J. Phys.: Condens. Matter*, Vol. 21, paper # 035111, 2009
- S. Keten and M.J. Buehler, "The strength limit of entropic elasticity in beta-sheet protein domains", *Physical Review E*, Vol. 78(6), paper number 061913, 2008

A selection of key technical publications are attached on the following pages

# Flaw Tolerance of Nuclear Intermediate Filament Lamina under Extreme Mechanical Deformation

Zhao Qin<sup>†</sup> and Markus J. Buehler<sup>†,‡,§,\*</sup>

<sup>†</sup>Laboratory for Atomistic and Molecular Mechanics (LAMM), Department of Civil and Environmental Engineering, <sup>‡</sup>Center for Computational Engineering, and <sup>§</sup>Center for Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

The nuclear lamina is a dense meshwork at the inner membrane of the nuclear envelope of eukaryotic cells<sup>1</sup> and is formed primarily by  $\sim$ 10 nm wide type-V intermediate filaments composed of lamin proteins.<sup>1–4</sup> The lamina contributes largely to the structural integrity of the nucleus, and therefore a key biological role is to protect genetic material against extreme environmental conditions.<sup>5–6</sup> The lamina is broken down during mitosis; this, however, represents only a relatively short period during the life of a cell.<sup>7,8</sup> Emphasizing the importance of this meshwork is the fact that there are at least 13 human diseases associated with mutations in lamin proteins. For example, the permanent loss of lamin proteins causes muscular dystrophy and cardiomyopathy, and overaccumulation of lamin mutants at the nuclear membrane causes the rapid aging disorder Hutchinson–Gilford progeria syndrome (HGPS).<sup>9–11</sup>

What is remarkable is that all the various types of nuclear lamina naturally feature structural imperfections, defects, and heterogeneities including those induced by nuclear pores.<sup>1,12–14</sup> Defects in the lamina can also be caused by localized forces from cytoskeletal structures or by atomic-scale defects due to the imperfect assembly of the meshwork, leading to void formation in the meshwork.<sup>1,12,13,15</sup> Another source of imperfections is generated by the clustering of lamin-associated nuclear pore complexes, which can grow rather large and are naturally present in virtually all cells.<sup>13,14,16</sup> Other possible factors include ionizing radiation that results in the breakage of covalent or non-covalent bonds within protein structures and thus results in structural defects.<sup>17–19</sup> It has been shown that continuous radiation, as experienced for example during space flight, causes cracks to occur around nuclear pores

**ABSTRACT** The nuclear lamina, composed of intermediate filaments, is a structural protein meshwork at the nuclear membrane that protects genetic material and regulates gene expression. Here we uncover the physical basis of the material design of nuclear lamina that enables it to withstand extreme mechanical deformation of  $>100\%$  strain despite the presence of structural defects. Through a simple *in silico* model we demonstrate that this is due to nanoscale mechanisms including protein unfolding, alpha-to-beta transition, and sliding, resulting in a characteristic nonlinear force–extension curve. At the larger microscale this leads to an extreme delocalization of mechanical energy dissipation, preventing catastrophic crack propagation. Yet, when catastrophic failure occurs under extreme loading, individual protein filaments are sacrificed rather than the entire meshwork. This mechanism is theoretically explained by a characteristic change of the tangent stress–strain hardening exponent under increasing strain. Our results elucidate the large extensibility of the nuclear lamina within muscle or skin tissue and potentially many other protein materials that are exposed to extreme mechanical conditions, and provide a new paradigm toward the *de novo* design of protein materials by engineering the nonlinear stress–strain response to facilitate flaw-tolerant behavior.

**KEYWORDS:** nuclear lamina · nuclear envelope breakdown · intermediate filament · biological material · multiscale modeling · flaw tolerance · materiomics

which leads to a complete fragmentation of nuclear envelopes once a critical defect concentration is reached<sup>17</sup> (see Figure S2 in the Supporting Information).

While experimental data has provided clear evidence that nuclei in healthy cells can effectively withstand extreme dilations and deformation without rupture,<sup>15</sup> conventional material models have failed to explain their significant capacity to expand without failure. This is because in most existing models locations of structural imperfections such as crack-like inclusions present singularities for stresses, which typically causes a deterioration of material properties. However, this is not observed in nuclear lamina which provides an extremely robust mechanical framework to protect genetic material.<sup>20</sup> Here we present an answer to this intriguing question and explain a general mechanism by which the intermediate filament meshwork, and

\* Address correspondence to mbuehler@mit.edu.

Received for review January 11, 2011 and accepted March 8, 2011.

Published online March 08, 2011

10.1021/mn200107u

© 2011 American Chemical Society

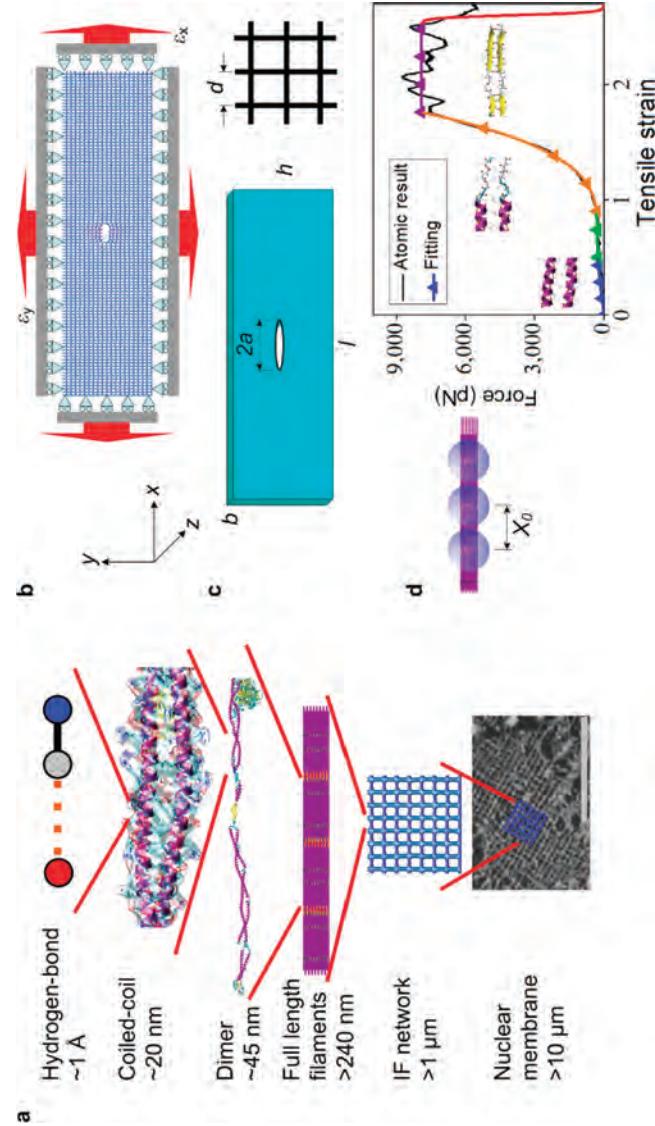
potentially other protein materials, are capable of forming an extremely robust flaw-tolerant material. We use a simple model developed based on a multi-scale approach to assess the mechanisms of defect initiation and propagation in the nuclear lamina under extreme loading conditions, mimicking the conditions a cell experiences during tissue stretching and contraction. The use of a multiscale approach to deal with this problem allows us to simulate a system with tens of micrometers dimension while retaining information about molecular structures and mechanisms.

## RESULTS AND DISCUSSION

**Computational Modeling.** We consider a meshwork model resembling a *Xenopus* oocyte nuclear lamina<sup>1,12</sup> because the structure and mechanical properties of this meshwork are experimentally well characterized, and it represents one of the few measured systems of a pure nuclear lamina without mechanical contributions of chromatin.<sup>13,15</sup> It is noted that so far only amphibian oocytes' nuclear lamina has been shown to form a rather regular lattice,<sup>1,13,15</sup> while the structure of nuclear lamina in other cells is likely more disordered and more tightly associated with chromatin.<sup>13</sup> Thus it is important to note that the model we study here may only represent a first approximation to the structure of other nuclear laminas. Figure 1a illustrates the hierarchical structure of this nuclear lamina.<sup>1,3,12,21</sup> We consider the meshwork with a small crack-like defect using uniaxial and biaxial tension (see Figure 1b, where the meshwork geometry is given in Figure 1c, with a characteristic defect size  $a$ ) to mimic *in situ* loading conditions. For instance, *in vivo* pressure differences across the nuclear envelope can result in an exposure of the nuclear lamina to significant biaxial tension.<sup>15</sup> As demonstrated by earlier experimental and computational molecular mechanics studies of the single intermediate filament level, the deformation mechanism is characterized first by elastic elongation,  $\alpha$  helix unfolding,<sup>22</sup> alpha-to-beta transition,<sup>23</sup> and eventual stick-slip-sliding<sup>24</sup> (Figure 1d; for details see Materials and Methods). This results in the characteristic nonlinear force-extension of a single fibril as shown in Figure 1d, with a severe change in the tangent modulus as strain is increased. As a model we use the mechanical response of vimentin intermediate filaments obtained in the earlier work<sup>21</sup> for exploring the mechanics of lamin meshwork. This is a reasonable first approximation because all intermediate filament proteins consist of tightly packed and aligned subunits, where each subunit is primarily composed of an extended coiled-coil structure which is crucial for the molecular mechanical properties as described above.<sup>25</sup> Indeed a direct comparison between lamin and vimentin molecular mechanics reveals that the mechanical response of the lamin dimer is found to be similar as the vimentin dimer (see Figure S4 in Supporting

Information). It has been shown in earlier works that this force-extension result agrees with experimental measurements for the general class of intermediate filament protein family<sup>21</sup> and we thus consider this in the formulation of our mesoscopic model to mimic the basic quasi-static deformation properties at vanishing pulling rates as relevant for physiological and experimental deformation speeds. This is an idealized mechanics model, which, although being limited by missing the diversity of intermediate filament types, offers a powerful starting point for understanding the fundamentals of the mechanical properties of intermediate filament meshworks. We emphasize that the goal of our model is not to accurately reflect the mechanical properties of one specific or all the different types of nuclear laminae. Rather it is formulated deliberately as a simple model to study generic aspects of the mechanical properties of protein meshwork materials at the microscale relating to the underlying molecular mechanics.

**The Flaw Tolerance of Nuclear Lamina.** On the basis of the knowledge of the behavior of a single fibril we now assess the mechanical response of an entire meshwork to large deformation by applying increasing uniaxial tension at the top and bottom layers. Figure 2a shows snapshots of the meshwork model with increasing deformation. We observe that no catastrophic failure of the meshwork occurs below 161% strain. However, we find that the crack geometry transforms drastically from an initial sharp edge in the  $x$ -direction into an elliptic shape, where the longest axis pointed in the  $y$ -direction, with a ratio between long axis and short axis being  $\sim 3$  (Figure 2a). What is remarkable is that instead of forming a rather small localized deformation zone at the crack tip as seen in linear elastic materials such as concrete, carbon fibers, or silica, virtually all filaments in the meshwork (oriented in the  $y$ -direction) undergo large deformation and unfold. This represents the formation of a very large "yield" region in which the elastic energy supplied by applied forces is dissipated. A detailed analysis of the forces that act on those filaments indeed confirms that the coiled-coil structure within each molecular subunit has undergone secondary structure unfolding *via* the breaking of clusters of hydrogen bonds up to strains of 90%. As the load is increased beyond 90% strain, a greater number of filaments begin to stiffen due to the onset of an alpha- $\beta$  transition at the molecular scale. For even greater loads, and when deformation reaches 161%, we observe that deformation starts to localize at the crack tip, resulting in eventual catastrophic failure of the meshwork. Under catastrophic failure the breakdown of the meshwork is mediated by highly localized breaking of individual filaments at the tip of the crack, and at the molecular scale due to slip of individual protein chains against each other (Figure 2a, snapshot iv).



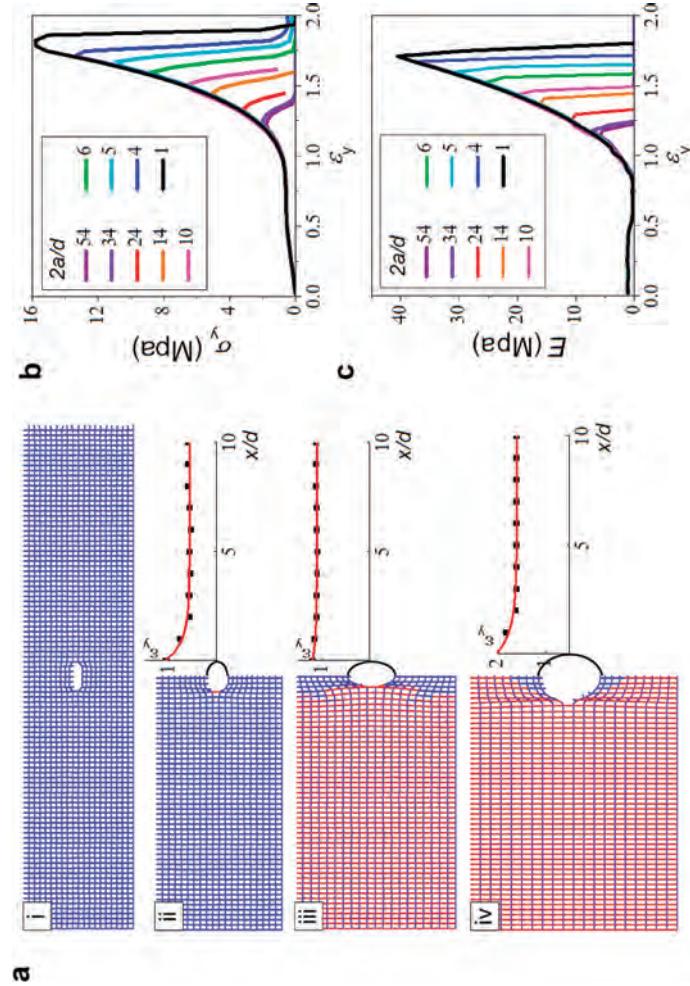
**Figure 1.** Hierarchical structure of nuclear lamina as found in the inner layer of nuclear envelop of *Xenopus* oocytes, boundary conditions in calculations, geometry of the meshwork model, and mechanical properties of a single intermediate filament. (a) Schematic of the hierarchical structure of nuclear lamina that ranges from nanoscale to macroscale. The image shows typical structural features of nuclear lamina, including hydrogen bonds, coiled-coil composed of  $\alpha$ -helices, a lamin A dimer composed of coiled-coil and linkers, bundles of dimers that fuse laterally to form full length filaments with a diameter of  $\sim 10$  nm, which form an orthogonal structural meshwork attached to the inner layer of nuclear envelope with a lattice constant of  $\sim 50$  nm $^{-1}$ . The scale bar in the nuclear membrane image is 1  $\mu$ m. (b) The orthogonal meshwork model boundary conditions, and the coordinate system used in our simulations. The initial crack is oriented in the  $x$ -direction and the meshwork is loaded in one direction (typically the  $y$ -direction) or in both directions. The consideration of different loading conditions mimics the randomness of the deformation field relative to crack directions in nuclear lamina. (c) The geometry parameters of the meshwork model. The parameters  $b$ ,  $l$ , and  $h$  are the thickness, length, and width of the meshwork model,  $d$  is the lattice constant ( $\sim 50$  nm) and  $2a$  is the initial crack length. (d) Schematic of the coarse-graining method with which each filament is modeled. A collection of mesoscopic beads are used to model the full length filament, where the equilibrated nearest-neighbor bead-to-bead distance is  $X_0$ , and a nonlinear interbead potential obtained from stretching tests of the full atomic intermediate filament model. No molecular unfolding is observed for  $\varepsilon < 50\%$ ,  $\alpha$  helix unfolding is observed as  $\varepsilon \geq 50\%$ , a alpha-to-beta transition is observed as  $\varepsilon \geq 90\%$ , and beta-sheets slide against each other beyond  $\varepsilon \geq 180\%$  and filaments fail at  $\varepsilon \geq 260\%$  strain. Inserted figures are snapshots of the atomic model of a segment of the 2B domain of the intermediate filament in stretching to visualize molecular mechanisms of deformation.

To summarize our observations, we find a sequence of events that involves delocalization of deformation at applied strains below 161% strain, followed by localization of deformation at larger applied strains and eventual catastrophic failure during which a small number of filaments break. The ability of the material to mitigate the effect of the defect by dissipating energy via unfolding of all filaments in the meshwork is referred to as flaw tolerance. This flaw tolerance can be theoretically explained by the characteristic nonlinear material behavior provided by the intermediate filament meshwork. The stress-strain curves for the bulk material (with and without defect) are depicted in Figure 2b and the Young's modulus ( $E$ ) vs strain ( $\varepsilon_y$ ) relation depicted in Figure 2c show that the meshwork features a highly nonlinear deformation character, with initial softening, followed by stiffening and eventual softening at the failure point. The lowest tangent modulus  $E$  is obtained at 70% strain and features a 3-fold decrease compared with the initial value at zero strain. The tangent modulus  $E$  then significantly increases with

strain, and at the failure point reaches a value that is 24-fold as large compared to its initial value.

**The Mechanism of Flaw Tolerance in the Nuclear Lamina.** The bulk strain–stress relation of a general nonlinear (that is, hyperelastic) material can be given by a simple power law as  $\sigma \sim \varepsilon^N$ , where  $N$  is the so-called hardening exponent that describes the stress–strain response of a material. Thereby,  $N < 1$  denotes softening behavior (also referred to as “elastic-plastic”), and  $N > 1$  represents a stiffening material. As seen directly in Figure 1d the behavior of the protein meshwork is therefore characterized by a change in the local hardening exponent  $N$  as the loading is increased, where for deformation up to 90%,  $N = 0$ , and for larger deformation of up to 176% strain,  $N \gg 1$ .

Why is this important? It is known from continuum fracture theory that crack-like defects such as voids typically result in stress and strain concentrations, which are mathematically described as singularities of stresses at the tip of the crack. The stress and strain singularity at the crack tip of a nonlinear material is



**Figure 2.** Deformation of the nuclear lamina under uniaxial loading at different levels of applied strain and mechanical analysis of the stress–strain properties of a meshwork with different crack sizes. (a) The snapshots i–iv show the deformation field for different levels of applied strain: i,  $\varepsilon = 38\%$ ; ii,  $\varepsilon = 58\%$ ; iii,  $\varepsilon = 110\%$ ; and iv,  $\varepsilon = 161\%$ . Each filament is colored according to its structural character, with blue denoting the structure before the alpha-to-beta transition has occurred, and red denoting the structure after the alpha-to-beta transition has occurred (at filament strains of around 90%). Tensile strains of filaments in the x-direction starting from the crack tip are plotted for snapshots ii–iv, where snapshot iv shows the onset of propagation of the crack. (b) Stress–strain relation for meshwork models with different initial crack lengths. (c) Tangent modulus-strain relation derived from panel b.

given by<sup>26</sup>

$$\frac{\sigma_c}{\sigma_f} = Fd^{N/(1+N)} \quad \text{and} \quad \frac{\varepsilon_c}{\varepsilon_f} = F^{1/N} d^{1/(1+N)} \quad (1)$$

where  $\sigma_c$  and  $\varepsilon_c$  represent the critical structural transition or failure stress and strain of a single filament at the crack tip,  $\sigma_f$  and  $\varepsilon_f$  are the applied stress and strain to the meshwork to reach  $\sigma_c$  and  $\varepsilon_c$  at the crack tip, and

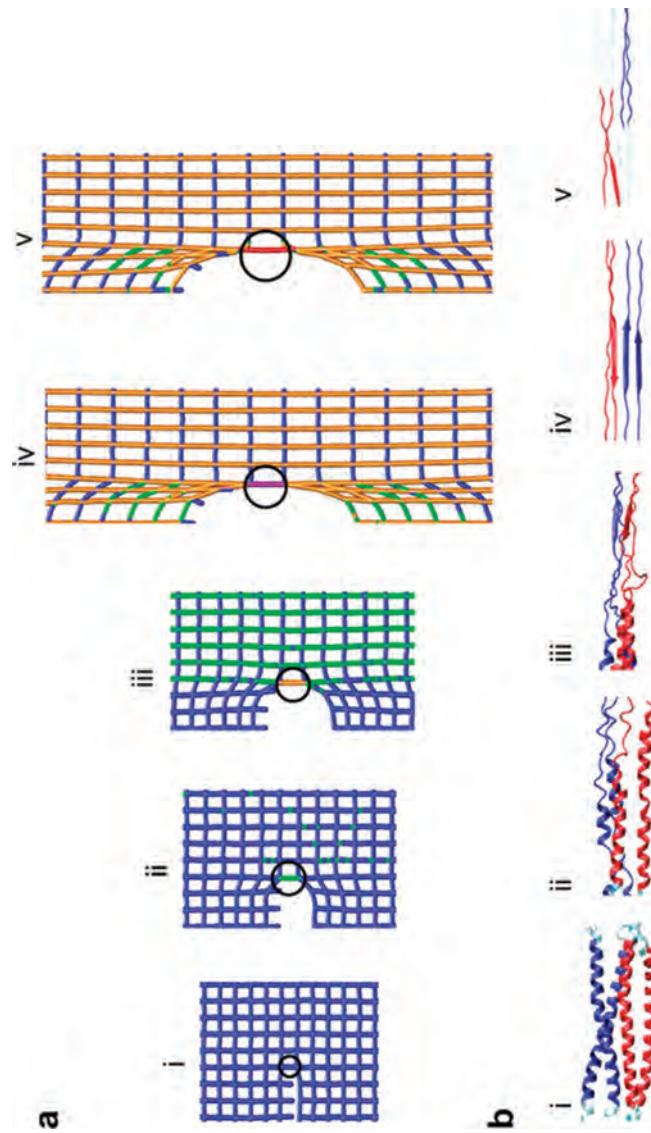
$F = [K(\sqrt{\pi}\sigma_c)]^{-[2N/(1+N)]}$  is a material constant ( $K$  is the material's fracture toughness).<sup>27</sup> The crack length,  $a$ , relates to size of the imperfections in the structure (see for example, in Figure 1c). Equation 1 shows that the nature of the stress or strain concentration depends strongly on the hardening exponent  $N$ .

We now use this simple model of how stresses and strains are distributed near a crack to explain the findings from our simulations. The first important insight used here is that the stress–strain response of the material in different regimes of deformation can be approximated by different values of  $N$ , reflecting the behavior seen in Figure 2b,c. Equation 1 predicts that a change in  $N$  leads to a drastic change in the stress and strain distribution near a crack tip. Specifically,  $N = 0$  leads to a highly delocalized stress field but a localized strain field, where  $N \gg 1$  leads to a highly localized stress field and a delocalized strain field. We note that here, “localized” means that a quantity is larger at the crack tip than elsewhere, and “delocalized” means that

a quantity is of similar magnitude at the crack tip than elsewhere in the system.

Indeed, in agreement with the notion that the hardening exponent  $N$  changes as the meshwork is deformed, we observe two distinct types of stress and strain distributions around the crack tip. First, in the softening regime of up to 90% strain, the strain of the highly stretched individual filament at the crack tip tends to reach the end of the softening stage of 90% and leaves filaments further away from the crack tip exposed to a much smaller (applied) strain of 58% (as shown in Figure 2a(iii)), and thus  $\varepsilon_c/\varepsilon_f = 1.55$ . This represents a localized distribution of strain, where all filaments oriented in the y-direction feature strains between those two bounds. However, because of the plateau in the force-strain curve the tensile forces that act on each filament for this range of strains are the same as shown in Figure 1d, implying that the stress is delocalized and  $\sigma_c/\sigma_f \approx 1$ . This corresponds to the situation when  $N \rightarrow 0$  (reflecting softening) in eq 1, which agrees with the hypothesis that this case leads to a delocalized stress distribution.

Second, in the stiffening stage from 90% up to 176% strain, we observe that the strain of filaments at the crack tip is 120%, and filaments further away from the crack tip are exposed to an applied strain of 110% (as shown in Figure 2a(iii)). While the strain distribution is delocalized, quantified by  $\varepsilon_c/\varepsilon_f \approx 1$ , it



**Figure 3.** Deformation of the lamin filaments around the crack tip under uniaxial loading at certain strain level. (a) Snapshots of the meshwork geometry near the crack tip at applied strains of (i)  $\varepsilon_y = 17\%$ , (ii) 42%, (iii) 58%, (iv) 160%, (v) 161%. The crack length is  $2a/b = 6$  and each filament is colored according to its structural character, as indicated in Figure 1d. (b) Visualization of the molecular structure of the intermediate filament under stretching. The structural character of the filament ahead of the crack tip (marked by a circle in panel a) is shown at different strains, corresponding to the strain levels shown in snapshots i–v in panel a (the local strain at the filament considered here is  $\varepsilon = 40\%$ , 50%, 90%, and 220% and thus generally larger than the applied strain).

leads to a highly localized stress distribution of  $\sigma_C/\sigma_F \approx 2.2$  (as shown in Figure 1d). We thus obtain  $N \gg 1$  (reflecting stiffening) from eq 1, which agrees with the hypothesis that this case leads to a localized stress distribution. In the large-deformation regime of over 161% strain and when rupture occurs, single filaments at the crack tip reach the maximum strength as shown in Figure 2a(iv) and break. This localized failure of single filaments at the crack tip results in eventual crack propagation (mediated by sequential breaking of single filaments at the crack tip) without any crack branching or other large-scale failure mechanisms that involve many filaments.

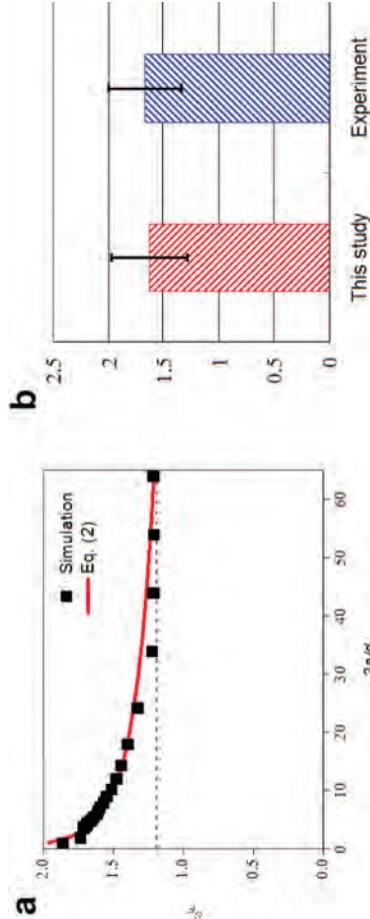
**Discussion of These Results.** Specific molecular mechanisms that act at different levels of strain control this change of the “local” value of  $N$  (in Figure 3a). Lamin filaments initially feature their typical structure as intact coiled-coils (Figure 3b), where clusters of hydrogen bonds are arranged in series and grouped in  $\alpha$ -helical turns. This geometry facilitates the unfolding of  $\alpha$ -helical loops after reaching the limit of elastic deformation at 50%.<sup>21</sup> The unfolding force until all  $\alpha$ -helical loops are unfolded is basically constant for long coils, leading to a strong softening behavior (which can be referred to as “ideal elastic-plastic”) and thus a hardening exponent of  $N=0$ . Further loading causes polypeptide strands to be fully unfolded, squeezed together driven by hydrophobic interactions and eventually undergoing a transition into beta-sheets. The

structure after this alpha-to-beta transition exhibits a stronger ability to withstand forces because many clusters of hydrogen bonds are loaded in parallel and deform cooperatively,<sup>23</sup> leading to a severe increase of the force and stiffness as loading is further increased. This results in severe stiffening and is captured in a hardening exponent of  $N \gg 1$ . An increasing loading causes first sliding and then rupture of the entire filament, leading to a final softening behavior at very large deformation. This analysis shows that the softening and stiffening behavior, described by changes in the “local” hardening exponent  $N$ , is caused by secondary structure transitions at the molecular level at the nanoscale.

Finally, in order to identify the flaw-tolerance ability of the material for varied defect sizes, we investigate the failure strain  $\varepsilon_F$  of the meshwork as a function of the crack length  $a$  as shown in Figure 4a. The results depicted in Figure 4a show that the failure strain is largely insensitive to the presence and size of cracks, and that the failure strain approaches a constant value even as the defect size grows. Considering eq 1 we empirically fit the  $\varepsilon_F-a$  relation by

$$\varepsilon_F = \varepsilon_{F0} \left( \frac{2a}{d} \right)^{-1/(1+N)} \quad (2)$$

where  $\varepsilon_{F0} = 196\%$  corresponds to the highest failure strain at  $2a = d$  (for the case where no crack is present, or a perfect meshwork), and  $N = 7.6$  is the hardening



**Figure 4.** Failure strain  $\varepsilon_f$  of the meshwork as a function of initial crack length, here quantified in normalized form (crack length normalized by the lattice constant  $d$ ). (a) Failure strain of the meshwork for varied crack sizes. The results show that the failure strain is largely insensitive to the presence and size of cracks, and that the failure strain approaches a constant value of 128% in the asymptotic limit for very large crack sizes. (b) The comparison of  $\varepsilon_f$  between results obtained in this study with a characteristic defect size and the experiment of the axial extensibility of nuclear lamina gel.<sup>8</sup> The data compares well and provides validation for the results obtained from our computational model.

exponent of nuclear lamina before failure. For the case of biaxial loading with  $\varepsilon_x = \varepsilon_y$ , we obtain  $\varepsilon_{F0} = 197\%$  and  $N = 6.8$ . It is noted that we confirmed that as we change the parameters  $h$ ,  $l$ , and  $d$  we find the parameters ( $\varepsilon_{F0}$ ,  $N$ ) do not change with those variations. The voids of nuclear lamina in experimental studies are observed to be up to on the order of  $1.4\text{ }\mu\text{m}$  (Supporting Information, Figure S3b), leading to  $\varepsilon_f = 128\%$ . This analysis explains the phenomenon that the integrity of nuclear lamina is not greatly affected by a single or larger number of clusters (for example, a conglomerate of several nuclear pores), representing a flaw-tolerant material.

Our results also explain recent experimental stretching tests of gels of nuclear lamina, where the material has been found to exhibit a large axial strain before failure in the range of 133%–200%.<sup>8</sup> This agrees with our finding that the failure strain of the nuclear lamina is between two bounds of 128%–197%, as shown in Figure 4b. Comparing with the failure strain measured for the entire nuclear envelope (in the range of 50%–60% strain<sup>15</sup>), the nuclear lamina is found to be much more extensible, and much greater than the deformation that can be generated by microtubules (<50%<sup>29</sup>) that impinge on the nuclear envelope.<sup>16</sup>

Therefore, external forces applied outside the cell are likely not sufficient to cause the rupture the nuclear lamina by mechanical signal transduction through microtubules. These results, together with the fact that the integrity of the nuclear lamina only breaks down during mitosis through breakdown processes, suggest that even if there exists a structural flaw, the nuclear lamina is a reliable structure that effectively protects genetic material from extreme external force and deformation applied on the cell, which will not result in rupture of the nuclear lamina unless there is further biochemical modification or mutations that affect its mechanical property to lead to a loss of flaw tolerance. Although the accuracy of this result may be distorted since the gel is composed of lamins in random

orientations, it is the only experimental result that quantitatively provides an estimate of the deformation capacity of nuclear lamina. We use this result to carry out the comparison here but are cognizant that future experimentation validations are required to fully understand all the mechanical properties of nuclear lamina. We note that other nuclear lamina models with different geometries have been proposed in other works.<sup>12,13,28</sup> Since both the *in situ* and *in vitro* experiments of somatic nuclear lamina are either not feasible or too difficult to carry out,<sup>13</sup> it is hard to tell which model is more immediately related to human somatic nuclear lamina. Since our model can in principle be readily adopted to different sets of geometries and mechanical parameters it may be used as a tool to determine which model is more realistic by comparing against future *in vivo* mechanical test in experiments. It is also noted that within our model we consider imperfections, defects, and heterogeneities simply as crack-like defects, which is a highly simplified representation since the biological model of defects can be more complex. For example, different models of the effect of the nuclear pore complex on the nuclear lamina have been discussed, and those models either consider a nuclear pore complex as a relatively weakly associated deformable unit similar to a void or strongly associated rigid complex similar to a rivet.<sup>28</sup> Although there is some experimental evidence that nuclear pore complexes act like rivets,<sup>28</sup> there is also evidence that rupture and breakdown of nuclear envelopes tend to occur around nuclear pore complexes.<sup>17</sup> The improvement of defect modeling of nuclear pore complexes requires more information of the interaction between the complex and nuclear lamina and has yet to be established.

## CONCLUSION

Using a simple model we identified here the physical basis of the material design of nuclear lamina that

enables it to withstand extreme mechanical deformation of more than 100% strain, despite the presence of structural defects. Our study shows that its integrity and strength are not compromised by structural defects such as nuclear pores or other imperfections embedded in nuclear envelopes and that this superior robustness is caused by mechanisms that act at multiple scales. This is facilitated by a combination of softening—stiffening—softening of filament mechanics, which is generated by protein unfolding, an alpha-to-beta transition, and intermolecular sliding under mechanical loading. The robust failure behavior of nuclear lamina is explained theoretically by a change of the stress—strain hardening exponent  $N$  as deformation is increased. We derived a general theoretical model that suggests that the critical feature of the dependence of  $N(\varepsilon)$  is that  $N \rightarrow 0$  (here, for strains up to 90%), and  $N \gg 1$  for larger strains (here, between 90% and 176%). This particular sequence of nonlinear stress—strain law of filaments facilitates an extremely delocalized deformation at strains of up to 176% and a highly localized failure once catastrophic breakdown of the material occurs at strains in excess of 176% strains. The delocalized deformation in the regime up to 176% vastly increases the energy absorbing capacity, while the localized failure during crack extension provides an important mechanism to ensure localized damage,

should it occur. Localized damage is critical for cells to be able to repair defects rather than having to deal with large-scale and widely distributed damage in the nuclear lamina. This suggests that the paradigm of dealing with failure, if it does occur, is that few protein filaments are sacrificed rather than the entire meshwork. As an important barrier to protect genetic material crucial to life, the nuclear lamina must indeed be an extremely reliable structure that can tolerate extreme mechanical conditions. The general mechanism to create such flaw-tolerant behavior revealed here, established through a particular nonlinear stress—strain response of constituting protein filaments realized through a series of molecular mechanisms, may also apply to other natural and synthetic materials, and could enable us to fabricate novel biologically inspired materials with high extensibility, high strength, robustness, and impact resistance.<sup>29</sup> This opens the possibility that by designing the molecular structure of the material the particular dependence of the hardening exponent  $N$  as a function of strain can be achieved. The multiscale approach used here could also enable studies of the mechanism of disease (e.g., rapid aging disease progeria), which is caused by point mutations at the protein level and that affects the mechanical properties of nuclear lamina by making it more brittle.

## MATERIALS AND METHODS

**Multiscale Modeling Method.** The mesoscopic model used here is set up based on a combination of experimental and atomistic data.<sup>1,12,21,30</sup> We use a simple mesoscopic model describing each intermediate filament as a series of beads interacting according to nonlinear interparticle multibody potentials. The model is designed by our desire to develop a simple model to derive generic insight into the mechanical properties and mechanisms. We note that even though such a simple model formulation does not allow us to derive quantitative conclusions for phenomena pertaining to specific types of nuclear lamina, it does enable us to understand universal, generic relationships between underlying molecular mechanisms, resulting nonlinear properties of the material, and the failure behavior of intermediate filament meshworks.

The total energy is given by

$$E_{\text{f}} = E_{\text{T}} + E_{\text{B}} \quad (3)$$

The total energy is given by the sum over all pair wise and three-body interactions:

$$E_{\text{T}} = \sum_{\text{pairs}} \varphi_{\text{T}}(r) \quad \text{and} \quad E_{\text{B}} = \sum_{\text{triplets}} \varphi_{\text{B}}(\theta) \quad (4)$$

Here we approximate the nonlinear force-extension behavior under tensile loading with a multipolynomial potential model that has been used successfully in the earlier study of deformation of cytoplasm single filament.<sup>31</sup> The tensile force between two particles (beads as illustrated in Supporting Information, Figure S1c) is described as

$$F_{\text{T}}(r) = -\frac{\partial \varphi_{\text{T}}(r)}{\partial r} \quad (5)$$

where

$$\frac{\partial \varphi_{\text{T}}(r)}{\partial r} = \left[ \exp\left(\frac{r - r_b}{r_b - \Xi}\right) + 1 \right]^{-1}$$

$$\begin{cases} k_1(r - r_0) & r < r_1 \\ R_1 + k_2(r - r_1) & r < r_2 \\ R_2 + k_3(r - r_2) + k_3^2(r - r_2)^2 + k_3^3(r - r_2)^3 & r < r_3 \\ R_3 & r \geq r_3 \end{cases} \quad (6)$$

In eq 6,  $k$  and  $r_i$  are spring constants that derived directly from the force-extension curve of the tension test of full atomic model (as shown in Supporting Information, Figure S1b, with their physical meaning and value defined in Supporting Information). The Fermi—Dirac distribution function introduces two additional parameters  $r_b$  and  $\Xi$ .<sup>32</sup> The parameter  $r_b$  denotes the critical separation distance for breaking of the filament and the parameter  $\Xi$  describes the amount of smoothing around the breaking point (the smaller  $\Xi$ , the smoother the curve becomes). A similar strategy to model the interatomic potential near rupture has been used in earlier work.

The bending energy (as illustrated in Supporting Information, Figure S1d) is given by

$$\varphi_{\text{B}}(\theta) = \frac{1}{2} k_{\text{B}}(\theta - \theta_0)^2 \quad (7)$$

with  $k_{\text{B}}$  relating to the bending stiffness of the intermediate filament  $E_{\text{f}}$  through  $k_{\text{B}} = 3E_{\text{f}}/r_0$ . The parameter  $E_{\text{f}}$  relates to the intermediate filament persistence length  $L_{\text{p}}$  through  $E_{\text{f}} = L_{\text{p}}k_{\text{B}}T$ , where  $k_{\text{B}}$  is the Boltzmann constant and  $T$  is the temperature.

The geometry of the meshwork is obtained from experimental observations of the nuclear lamina of *Xenopus* oocytes.<sup>1</sup> The meshwork is composed of two sets of near-orthogonal

intermediate filaments with a lattice constant of  $d = 50$  nm, and we use  $d$  to normalize the crack length, and that value gives the size of defects. We use mesoscopic beads to model the each intermediate filament within the meshwork and use  $r_0 = 5$  nm, this length is much smaller than the persistence length of the full length filament ( $\sim 1$   $\mu$ m), and it equals to the radius of a full length filament.

**Computational Experiments.** Calculations are carried out in two steps, first relaxation followed by loading. Relaxation is achieved by heating up the system, then annealing the structure at temperature of 300 K, followed by energy minimization. After relaxation, we keep the system at 300 K in an NVT ensemble (constant temperature, constant volume, and constant number of particles) and apply loading by displacement boundary conditions (by fixing one single layer of beads near the meshwork boundary and apply the load in normal directions. This setup resembles linear rails with guides clamped to the boundary, resulting in biaxial deformation independently<sup>33</sup>), continuously displacing particles in the boundary in a speed of 0.01  $\text{Å}/\text{ps}$ . Studies with varying loading rates are carried out based on the mesoscopic model of single filaments. By comparing with force-extension curves obtained from atomic model of vimentin intermediate filaments and experimental studies of several classes of intermediate filaments near equilibrium<sup>21</sup>, it is confirmed that the molecular model is validated against experimental data.

**Damping.** Damping effects are included in the model by considering the energy dissipation of the intermediate filament in motion caused by the viscosity of water environment. We note because of the small character dimension of the intermediate filament, the Reynolds number is  $\ll 2300^{34}$  and the drag force is approximately proportional to particle velocity (laminar flow). We use the Stokes' law to measure the drag force by<sup>35</sup>

$$f_{\text{drag}} = -6\pi\eta Rv \quad (8)$$

where  $\eta = 8.6 \times 10^{-4} \text{ Pa}\cdot\text{s}$  is the fluid viscosity constant of water at room temperature, and  $R$  is the equivalent spherical radius of the mesoscopic bead, which equals the radius of a sphere of equivalent volume given by<sup>35</sup>

$$R = \left( \frac{3}{16} b^2 r_0 \right)^{1/3} \quad (9)$$

with geometry parameters of the mesoscopic bead given as  $R = 4.5$  nm. The parameter  $v$  denotes the relative velocity of particle motion in a continuous viscous fluid.

**Acknowledgment.** We acknowledge discussions with Kris Dahl (Carnegie Mellon). Support for this research was provided by AFOSR. **Supporting Information Available:** Additional information regarding the computational model development, validation, and computing technique and supporting references. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES AND NOTES

- Aebi, U.; Cohn, J.; Buhle, L.; Gerace, L. The Nuclear Lamina Is a Meshwork of Intermediate-Type Filaments. *Nature* **1986**, *323*, 560–564.
- Chang, L.; Shav-Tal, Y.; Tricek, T.; Singer, R. H.; Goldman, R. D. Assembling an Intermediate Filament by Dynamic Cotranslation. *J. Cell Biol.* **2008**, *177*, 747–756.
- Hermann, H.; Bar, H.; Kieplak, L.; Strelkov, S. V.; Aebi, U. Intermediate Filaments: From Cell Architecture to Nanomechanics. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 562–573.
- Ishikawa, H.; Bischoff, R.; Holtzer, H. Mitosis and Intermediate-Sized Filaments in Developing Skeletal Muscle. *J. Cell Biol.* **1968**, *38*, 538–555.
- Burke, B.; Stewart, C. L. Life at the Edge: The Nuclear Envelope and Human Disease. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 575–585.
- Qin, Z.; Buehler, M. J.; Kieplak, L. A Multiscale Approach to Understand the Mechanobiology of Intermediate Filaments. *J. Biomech.* **2010**, *43*, 15–22.
- Guttinger, S.; Laurell, E.; Kutay, U. Orchestrating Nuclear Envelope Disassembly and Reassembly During Mitosis. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 178–191.
- Panorchan, P.; Schafer, B. W.; Witz, D.; Tseng, Y. Nuclear Envelope Breakdown Requires Overcoming the Mechanical Integrity of the Nuclear Lamina. *J. Biol. Chem.* **2004**, *279*, 43462–43467.
- Eriksson, M.; Brown, W. T.; Gordon, L. B.; Glynn, M. W.; Singer, J.; Scott, L.; Erdos, M. R.; Robbins, C. M.; Moses, T. Y.; Berglund, J.; et al. Recurrent *de Novo* Point Mutations in Lamin A Cause Hutchinson–Gilford Progeria Syndrome. *Nature* **2003**, *423*, 293–296.
- Omari, M. B.; Coulombe, P. A.; McLean, W. H. I. Mechanisms of Disease: Intermediate Filament Proteins and Their Associated Diseases. *New Engl. J. Med.* **2004**, *351*, 2087–2100.
- van der Kooi, A. J.; Bonne, G.; Eymard, B.; Duboc, D.; Tilmis, L.; et al. Lamin a/C Mutations with Lipodystrophy, Cardiac Abnormalities, and Muscular Dystrophy. *Neurology* **2002**, *59*, 620–623.
- Goldberg, M. W.; Fiserova, J.; Huttengauch, J.; Stick, R. A. New Model for Nuclear Lamina Organization. *Biochem. Soc. Trans.* **2008**, *36*, 1339–1343.
- Goldberg, M. W.; Huttengauch, J.; Hutchison, C. J.; Stick, R. Filaments Made from a- and B-Type Lamins Differ in Structure and Organization. *J. Cell Sci.* **2008**, *121*, 215–225.
- Schemelieh, L.; Carlton, P. M.; Haase, S.; Shao, L.; Winoto, L.; Kner, P.; Burke, B.; Cardoso, M. C.; Agard, D. A.; Gustafsson, M. G.; Leonhardt, H.; Sedat, J. W. Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3d Structured Illumination Microscopy. *Science* **2008**, *320*, 1332–1336.
- Dahl, K. N.; Kahn, S. M.; Wilson, K. L.; Discher, D. E. The Nuclear Envelope Lamina Network Has Elasticity and a Compressibility Limit Suggestive of a Molecular Shock Absorber. *J. Cell Sci.* **2004**, *177*, 4779–4786.
- Beaudouin, J.; Gerlich, D.; Daigle, N.; Ellis, R.; Ellenberg, J. Nuclear Envelope Breakdown Proceeds by Microtubule-Induced Tearing of the Lamina. *Cell* **2002**, *108*, 83–96.
- Diazcenz, J. W.; Starzyk, H.; Ruzicchio, Z. X-ray-Irradiation Induced Changes of Nuclear Membrane of Kirkman–Robbins Tumor-Cells. *Experientia* **1973**, *29*, 83–84.
- Munro, T. R. Relative Radiosensitivity of Nucleus and Cytoplasm of Chinese Hamster Fibroblasts. *Radiat. Res.* **1970**, *42*, 451–470.
- Székely, J. G.; Copps, T. P.; Morash, B. D. Radiation-Induced Invagination of the Nuclear-Envelope. *Radiat. Res.* **1980**, *83*, 621–632.
- Jamay, P. A.; Euteneuer, U.; Traub, P.; Schliwa, M. Viscoelastic Properties of Vimentin Compared with Other Filamentous Biopolymer Networks. *J. Cell Biol.* **1991**, *113*, 155–160.
- Qin, Z.; Kieplak, L.; Buehler, M. J. Hierarchical Structure Controls Nanomechanical Properties of Vimentin Intermediate Filaments. *PLOS ONE* **2009**, *4*, e7294.
- Ackbarow, T.; Chen, X.; Keten, S.; Buehler, M. J. Hierarchies, Multiple Energy Barriers, and Robustness Govern the Fracture Mechanics of Alpha-Helical and Beta-Sheet Protein Domains. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16410–16415.
- Qin, Z.; Buehler, M. J. Molecular Dynamics Simulation of the Alpha-Helix to Beta-Sheet Transition in Coiled Protein Filaments: Insights from Hagfish Slime Threads. *Biophys. J.* **2003**, *85*, 2015–2027.
- Rice, J. R.; Rosengren, G. F. Plane Strain Deformation near a Crack Tip in a Power-Law Hardening Material. *J. Mech. Phys. Solids* **1963**, *16*, 1–12.
- Pugno, N.; Carpinieri, A.; Ippolito, M.; Mattoni, A.; Colombo, L. Atomistic Fracture: QFM vs MD. *Eng. Fract. Mech.* **2008**, *75*, 1794–1803.

28. Rowat, A. C.; Foster, L. J.; Nielsen, M. M.; Weiss, M.; Ipsen, J. H. Characterization of the Elastic Properties of the Nuclear Envelope. *J. R. Soc. Interface* **2005**, *2*, 63–69.
29. Buehler, M. J.; Yung, Y. C. Deformation and Failure of Protein Materials in Physiologically Extreme Conditions and Disease. *Nat. Mater.* **2009**, *8*, 175–188.
30. Xu, Z. P.; Paparcone, R.; Buehler, M. J. Alzheimer's  $\beta$ -Amyloid Fibrils Feature Size-Dependent Mechanical Properties. *Biophys. J.* **2010**, *98*, 2053–2062.
31. Beraud, J.; Qin, Z.; Buehler, M. J. Intermediate Filament-Deficient Cells Are Mechanically Softer at Large Deformation: A Multiscale Simulation Study. *Acta Biomater.* **2010**, *6*, 2457–2466.
32. Buehler, M. J.; Gao, H. J. Dynamical Fracture Instabilities Due to Local Hyperelasticity at Crack Tips. *Nature* **2006**, *439*, 307–310.
33. Petersan, P. J.; Deegan, R. D.; Marder, M.; Swinney, H. L. Cracks in Rubber under Tension Exceed the Shear Wave Speed. *Phys. Rev. Lett.* **2004**, *93*, 015504.
34. Batchelor, G. K. *An Introduction to Fluid Dynamics*; 1st Cambridge Mathematical Library ed.; Cambridge University Press: Cambridge, U.K.; New York, 2000.
35. Jennings, B. R.; Parslow, K. Particle-Size Measurement—The Equivalent Spherical Diameter. *Proc. R. Soc. London, Ser. A* **1988**, *419*, 137–149.

# Dynamic Failure of a Lamina Meshwork in Cell Nuclei under Extreme Mechanical Deformation

Zhao Qin · Markus J. Buehler

© Springer Science+Business Media, LLC 2011

**Abstract** The nuclear lamina is a structural protein meshwork at the inner nuclear membrane. It confers mechanical strength to the cell's nucleus and also sustains the overall structural integrity of the cell. The rupture of nuclear lamina is involved in many physiologically extreme conditions, such as cell division, genetic disease, and injury. Yet, its rupture mechanisms and processes are largely unknown and failure models commonly used for engineering materials cannot be directly applied due to the complex hierarchical structure. Here, we use a multiscale modeling technique to investigate the dynamic failure of the nuclear lamina meshwork from the bottom up. We find that flaws or cracks in the nuclear lamina act as seeds for catastrophic failure that propagate rapidly upon very large deformation. Fracture occurs via crack propagation at interersonic speeds, and greater than the Rayleigh-wave speed predicted as a limit by classical fracture theory but smaller than the longitudinal wave speed. Our analysis shows that nanoscale secondary structural changes in protein filaments such as the alpha–beta transition and intermolecular sliding explain this macroscale phenomenon. Based on a simple model, we discover that the crack propagation speed is governed by the square root of the ratio of the tangent material moduli in ( $E_x$ ) and perpendicular ( $E_y$ ) to the crack propagation direction,  $v \sim \sqrt{E_x/E_y}$  where the relative levels of applied strains in the  $x$ - and  $y$ -direction control the crack speed.

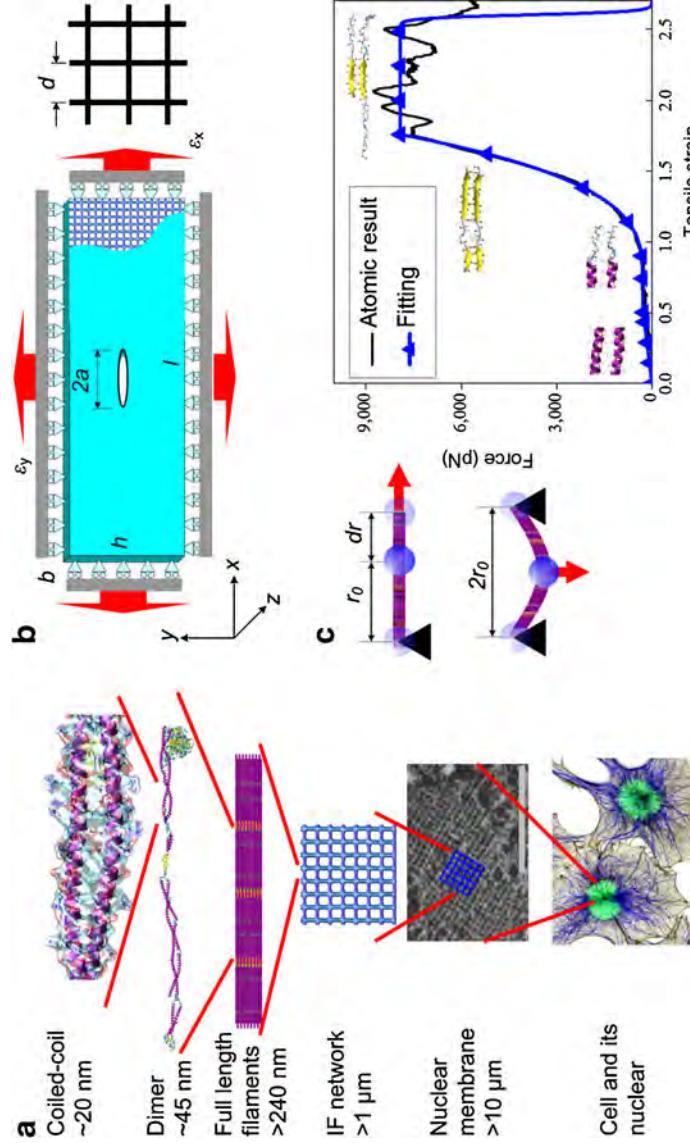
**Keywords** Nuclear lamina · Intermediate filament meshwork · Biological material · Multi-scale modeling · Failure · Materionics

## 1 Introduction

The term intermediate filament refers to a heterogeneous family of cytoskeletal and nuclear skeletal proteins with the ability to self-assemble into  $\approx$ 10-nm-wide filaments [1–4]. The name was coined back in 1968 because their diameters appeared to be intermediate in size between those of two other prominent cellular components, microtubules and microfilaments [3]. Intermediate filaments are fibrous proteins that are absent from both plants and fungi, but found widely in animals; and they are linked to several human diseases [5–8]. Lamin is a type of intermediate filament that is found in the nuclear lamina meshwork, which is a key building block that maintains the integrity of the cell's nuclear envelope [2]. The nuclear lamina features high strength and extensibility, enabling it to protect the genetic material inside the nucleus from damage. However, physiologically extreme conditions, including contraction of muscular cells, depolymerization during cell division, the exposition to ionizing radiation, and genetic mutations, can lead to flaws and defects within the meshwork structure. Accumulated flaws can result in the generation of crack-like defects as shown in earlier experimental work [9].

Much attention has been paid to the physiological characteristics of the nuclear lamina, and especially to the factors that affect the stability of the nuclear envelope. The nuclear lamina is found to be highly mutation sensitive, and a complex set of more than 13 different human diseases including Hutchinson–Gilford progeria syndrome (rapid-aging disease) [5, 6, 8] is found in humans. For example, the rapid-aging disease originates from a point mutation at the tail domain of the lamin

Z. Qin · M. J. Buehler (✉)  
Laboratory for Atomistic and Molecular Mechanics,  
Department of Civil and Environmental Engineering,  
Massachusetts Institute of Technology,  
77 Mass. Ave. Room 1-235A&B,  
Cambridge, MA 02139, USA  
e-mail: mbuehler@MIT.EDU



**Fig. 1** Hierarchical structure of an intermediate filament meshwork as found in the inner layer of nuclear envelope of *Xenopus* oocytes, boundary conditions in simulations, geometry of the meshwork model and tensile properties of a single intermediate filament. **a** Schematic of the hierarchical intermediate filament structure that ranges from nano to macro. Panel **a** shows typical structural features of intermediate filament meshwork, including hydrogen bonds, coiled-coil composed of alpha-helix, a lamin A dimer composed of coiled-coil and linkers, bundles of dimers that fuse laterally to form full length filaments with a diameter of  $b \approx 10$  nm, which form an orthogonal meshwork attached to the inner nuclear membrane with a lattice constant of  $\approx 50$  nm [2]. The scale bar is 1  $\mu\text{m}$ . **b** The geometry of the orthogonal meshwork model, boundary conditions, and the coordinate system used in our

subunit, and is also known to alter the mechanical properties of the entire nuclear envelope [1, 6, 8, 10, 11]. Recent research revealed that defect accumulation in lamin also plays an important role in normal aging of human beings [12]. It has been shown in experimental studies that mutations and defects in nuclear lamina affect the nuclear envelope breakdown process during cell division, which can be a direct reason to cause illness. However, no studies have thus far linked the physical properties of the nuclear lamina to its underlying atomic and molecular structure and provided an explanation for its fundamental failure mechanism across multiple scales as has been done for other (e.g. engineered) materials [13–15].

In this paper, we investigate the mechanical failure of the nuclear lamina meshwork. We use a simple coarse-grain meshwork model with a realistic geometric and physical properties derived directly from the molecular scale to mimic the lamina meshwork of *Xenopus* oocytes [2, 16] as shown in Fig. 1a. We symmetrically fix one single layer of coarse-grain particles (so-called “beads”) near the meshwork boundary and apply loading as shown in Fig. 1b. This

simulations. The initial crack is oriented into the  $x$ -direction and the meshwork is loaded in one or both directions. The parameters  $b$ ,  $l$ , and  $h$  denote the thickness, length, and width of the meshwork model;  $d$  is the lattice constant ( $\approx 50$  nm) and  $2a$  is the initial crack length. **c** Schematic of the coarse-grain method used to model the meshwork. Mesoscopic particles, or “beads” are used to describe full length filaments, where the equilibrated nearest neighbor distance is  $r_0$  and an interbead potential fitted from stretching tests based on full atomistic modeling of intermediate filament is adopted. Inserted figures are snapshots of the atomic model of a segment of the 2B domain of an intermediate filament dimer under stretching, revealing molecular mechanisms of deformation

setup resembles linear rails with guides clamped to the boundary as done in earlier experimental and computational studies [17, 18]. The mechanical response of each filament in the meshwork is modeled according to the mechanical response of a corresponding full atomistic model under tension. The deformation mechanisms revealed by a series of earlier atomistic modeling [19, 20] include: (a) a yielding regime (unfolding of alpha-helices), (b) a stiffening regime (alpha-to-beta structural transition), and (c) a secondary yielding regime (interprotein sliding under breaking of interprotein beta-strands). These fundamental mechanisms of deformation of each filament under loading are summarized in Fig. 1c.

## 2 Materials and Methods

### 2.1 Simulation Setup

This study uses a coarse-grain approach to implement a mechanical model of the intermediate filament meshwork

based on a hierarchical finer-triangular-coarser multi-scale scheme. The entire structure of the nuclear envelope involves many other components, including the chromatin (resembling amongst others the genetic material) and cytoskeletal proteins. So far, however, only the amphibian oocytes' nuclear lamina has been isolated in experiments without association with chromatin [21]. Therefore the model we study here represents a first approximation to the realistic structure of other nuclear lamina, which may be more disordered and/or have slightly variegated structures. We also note that it differs from the physiological process of nuclear envelope breakdown process during cells' division since other chemical effects on the meshwork are not included in our work. This is motivated by the fact that it has been shown that an important function of the nuclear lamina is maintaining the mechanical integrity of cell nuclei, which is the focal point of this work [22].

The meshwork model shown in Fig. 1b includes 20 filaments in the  $y$ -direction ( $1 \mu\text{m}$ ) and 100 filaments ( $5 \mu\text{m}$ ) in the  $x$ -direction, with a crack-like defect added in the center to mimic a structural imperfection [9]. Filaments at the orthogonal corners are cross-linked by strong, covalent bonds which cannot break apart in our model. This assumption is based on the fact that the covalent bonds forming these cross-links are much stronger than the forces needed to cause filament rupture by sliding. All simulations reported here are carried out in two steps. First, we perform a relaxation during which we equilibrate the system. Relaxation is achieved by energy minimization, heating up the system from 0 to 300 K, then annealing the structure at a temperature of 300 K. Second, we perform a loading simulation during which we keep the system at 300 K using an  $NVT$  ensemble and apply a constant strain rate of  $0.13 \text{ ns}^{-1}$  to continuously increase the loading applied until the crack begins to travel through the entire meshwork. In studying the crack speed-to-applied strain relationships, we

repeat our simulations, stop the increasing loading, fix the boundary at target strains, and record the steady-state crack propagation speeds.

## 2.2 Multi-scale Modeling Approach

The mesoscopic model is set up based on a combination of experimental and full atomic calculation data [2, 19]. We use a mesoscopic model describing each intermediate filament as a series of beads interacting according to nonlinear interparticle multibody potentials. The total energy is given by:

$$E_x = E_T + E_B, \quad (1)$$

The total energy is given by the sum over all pair-wise and three-body interactions:

$$E_T = \sum_{\text{pair}} \varphi_T(r), E_B = \sum_{\text{triplets}} \varphi_B(\theta) \quad (2)$$

where  $\varphi_T(r)$  defines the pair-wise interaction potential between two bonded beads as a function of the distance  $r$  between the two beads, and  $\varphi_B(\theta)$  defines the three-body interaction potential among three bonded beads as a function of the interior angle  $\theta$  of the two pair bonds. Here, we approximate the nonlinear force-extension behavior under tensile loading with a multi-polynomial potential model that has been used successfully in an earlier study of deformation of single filaments found in the cytoplasm [23]. The tensile force between two particles (beads as illustrated in Fig. 1c) is described as:

$$F_T(r) = -\partial \varphi_T(r) / \partial r, \quad (3)$$

Where

$$\frac{\partial \varphi_T(r)}{\partial r} = [\exp\left(\frac{r - r_b}{r_b} \Xi\right) + 1]^{-1} \begin{cases} k_1(r - r_0) & r < r_1 \\ R_1 + k_2(r - r_1) & r < r_2 \\ R_2 + k_3^1(r - r_2) + k_3^2(r - r_2)^2 + k_3^3(r - r_2)^3 & r < r_3 \\ R_3 & r \geq r_3 \end{cases} \quad (4)$$

In Eq. 5,  $k_i$  and  $r_i$  are spring constants that derived directly from the force-extension curve of the tension test of full atomic model (as shown in Fig. 1c).  $r_0$  is the equilibrium distance between pair beads. Alpha-helix domains within the filament are intact before the extension reaches  $r_1$  and the slope of the force-extension curve is constant as defined by  $k_1$ . The extension beyond  $r_1$  leads to the unfolding of the alpha-helix domains until they become nearly fully unfolded at  $r_2$ , and in this extension region the

force-extension curve is linear with a slope of  $k_2$ . The third region before the extension of  $r_3$  corresponds to the stiffening of the material because of the alpha–beta transition, and the nonlinearity of the force extension curve is expressed by a polynomial function with stiffness parameters  $k_3^1$ ,  $k_3^2$ , and  $k_3^3$ . In the fourth region, the subunits within the filament slide against each other under a constant force. The parameters  $R_1$ ,  $R_2$ , and  $R_3$  are calculated from force continuity conditions.

**Table 1** Geometric and numerical parameters for the coarse-grained computational model

Parameter and units	Numerical value
Equilibrium bead distance $r_0$ (in Å)	50
Critical distances $r_1, r_2, r_3$ (in Å)	75, 95, 138, 180
Tensile stiffness parameters $k_1, k_2$ (in kcal/mol/Å <sup>2</sup> )	0.1595, 0.0324,
Tensile stiffness parameters for nonlinear region $k_1^3$ (in kcal/mol/Å <sup>2</sup> ), $k_2^3$ (in kcal/mol/Å <sup>3</sup> ), $k_3^3$ (in kcal/mol/Å <sup>4</sup> )	0.2044, 0.0146, $9.2465 \times 10^{-4}$
Force continuity conditions $R_1, R_2$ , and $R_3$ (in kcal/mol/Å)	3.9877, 4.6357, 113.9364
Bond breaking distance $r_b$ (in Å)	180
Equilibrium angle $\theta_0$ (in rad)	$\pi$
Bending stiffness parameter $r_b$ (in kcal/mol/rad <sup>2</sup> )	169.51
Mass of each mesoscale particle (in amu)	230913
Switch function $\Xi$	300
Density $\rho$ (in kg/m <sup>3</sup> )	260

The value of these parameters are given in Table 1 without empirical fitting and are based directly on molecular modeling results. The Fermi–Dirac distribution function introduces two additional parameters  $r_b$  and  $\Xi$  [24]. The parameter  $r_b$  denotes the critical separation distance for breaking of the filament and the parameter  $\Xi$  describes the amount of smoothing around the breaking point (the smaller  $\Xi$ , the smoother the curve becomes). A similar strategy to model the potential near rupture for numerical analyses has been used in earlier studies. The numerical value of  $\Xi$  is set to a large value ( $\Xi=300$ ) similar as in earlier work [24], and is chosen to mimic the steep descent characteristics of the force-extension curve as obtained from the molecular-level tetrameter model as depicted in Fig. 1c.

The bending energy is given by:

$$\varphi_B(\theta) = \frac{1}{2} k_B (\theta - \theta_0)^2 \quad (5)$$

with  $k_B$  relating to the bending stiffness of the intermediate filament  $EI$  through  $k_B = 3EI/\theta_0$ . The parameter  $EI$  relates to the intermediate filament persistence length  $L_p$  through  $EI=L_p K_B T$ , where  $K_B$  is the Boltzmann constant and  $T$  is the temperature.

The geometry of the meshwork is obtained from experimental observations of the nuclear lamina of *Xenopus* oocytes [2]. The meshwork is composed of two sets of near-orthogonal intermediate filaments with a lattice constant of  $d=50$  nm, and we use  $d$  to normalize the crack length. We use a mesoscopic beads to model the each intermediate filament within the meshwork and define  $r_0=5$  nm; this length is much smaller than the persistence length of the full length filament (on the order of 1 μm) and it is equal to the radius of the full length filament.

### 2.3 Damping

The damping effect is included in the model by considering the energy dissipation of the intermediate filament in motion caused by the viscosity of water environment. We note because of the small characteristic dimension of intermediate filaments, the Reynolds number  $\ll 2,300$  [25] and the drag force is approximately proportional to particle velocity (laminar flow). We use Stokes' law to estimate the drag force by [25]:

$$f_{\text{drag}} = -6\pi\mu Rv, \quad (6)$$

where  $\mu = 8.6 \times 10^{-4}$  Pa · s is the fluid viscosity constant of water at room temperature,  $R$  is the equivalent spherical radius of the mesoscopic bead, which equals to the radius of a sphere of equivalent volume given by [26]

$$R = \left( \frac{3}{16} b^2 r_0 \right)^{1/3} \quad (7)$$

with the geometry parameters of the mesoscopic bead (here,  $b=10$  nm is the averaged diameter of intermediate filaments as measured in experimental work [1–4]), and we have  $R=4.5$  nm. The parameter  $v$  denotes the relative velocity of the particle motion in a continuous viscous fluid.

### 2.4 Crack Propagation Velocity Measurements

The rupture length for bead–bead bonds is defined as  $1.03r_b$ , where the corresponding value of the Fermi–Dirac function is  $<5 \times 10^{-5}$ . The crack tip position is determined by finding the  $x$ -coordinate of the interior surface particle with maximum  $x$  or minimum  $x$  positions. We obtain the crack length by measuring the length between the two crack tips and average it over a small time interval to eliminate high-frequency fluctuations. We differentiate the half length with respect to time to obtain the crack propagation velocity for each of the two crack tips. To obtain the steady-state velocity of the crack, we ensure that the crack onset is within the region that the crack length satisfies  $2L/d < 35$ ; and we average the crack propagation speed as it reaches C beyond this region.

### 2.5 Computing Technique

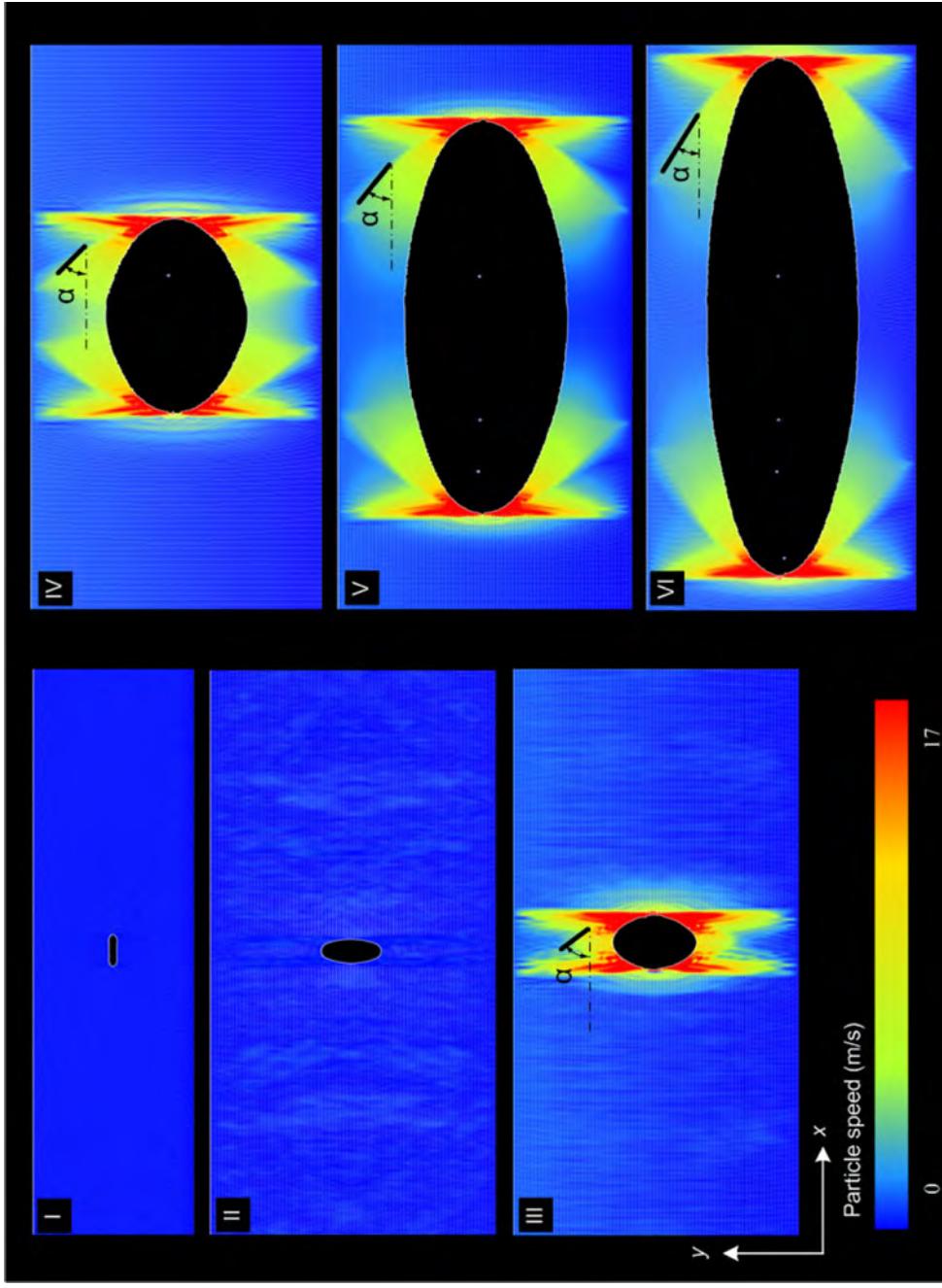
The molecular simulations described here are carried out with a classical molecular dynamics approach by using an extended version of the LAMMPS code [27]. Visualization of atomic and mesoscopic models is carried out with the Visual Molecular Dynamics (VMD) software. Hydrogen bonds are defined to be within 5 Å for visualization purposes for atomic models for the inlays shown in Fig. 1 [28]. Visualization of the meshwork model is performed using MATLAB.

### 3 Results and Discussion

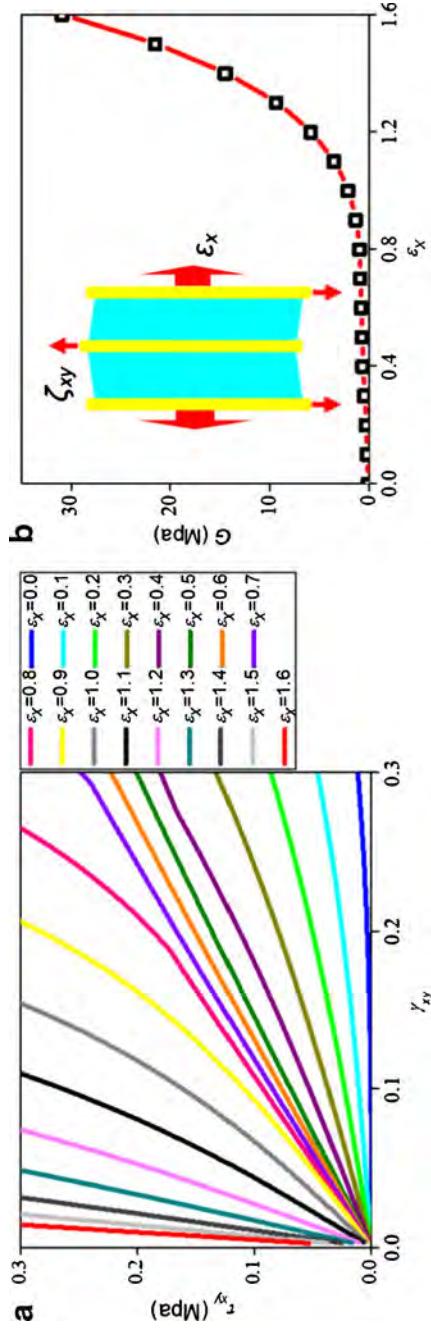
Figure 2 illustrates the failure mechanism of the meshwork with several snapshots shown for the case of uniaxial tensile crack loading (resembling mode I). To realize uniaxial tensile loading only  $\varepsilon_y$  is increased with a constant lateral strain  $\varepsilon_x=0$ . The meshwork is deformed and snapshots are taken at applied strains of  $\varepsilon_{y1}=37\%$ ,  $\varepsilon_{y2}=161\%$ ,  $\varepsilon_{y3}=162\%$ ,  $\varepsilon_{y4}=167\%$ ,  $\varepsilon_{y5}=172\%$ , and  $\varepsilon_{y6}=174\%$ . We observe no failure before  $\varepsilon_{y2}=161\%$ , but the crack transforms from an initial horizontal orientation with a sharp edge pointed in the  $x$ -direction into a vertical orientation where the longest axis points in the  $y$ -direction. We observe that instead of forming a highly localized yield zone at the crack tip as typically seen in materials with a crack, all filaments in the direction in which tensile load is applied stretch significantly, and deformation extends over the entire specimen and is not limited to the crack tip region.

A detailed analysis of the molecular state of these filaments reveals that the coiled-coil structure within each dimer building block has undergone molecular unfolding and a subsequent alpha–beta transition [20]. Catastrophic failure of the meshwork occurs once loading reaches  $\varepsilon_{y3}=162\%$  and proceeds by rapid extension of the initial crack, which continues to propagate as a rapidly moving single crack as is shown in Fig. 2 (III–VI). At the single filament scale, crack propagation is facilitated by the sliding apart of protein filaments after the transition into antiparallel beta-strands is completed, as is shown in Fig. 1c. Crack propagation occurs in a straight fashion into the  $x$ -direction without any branching or other instabilities.

We proceed with a detailed analysis of the dynamics of failure by crack propagation. We measure the steady-state crack speed  $C$  and find that  $C=18$  m/s,  $C=21$  m/s,  $C=$



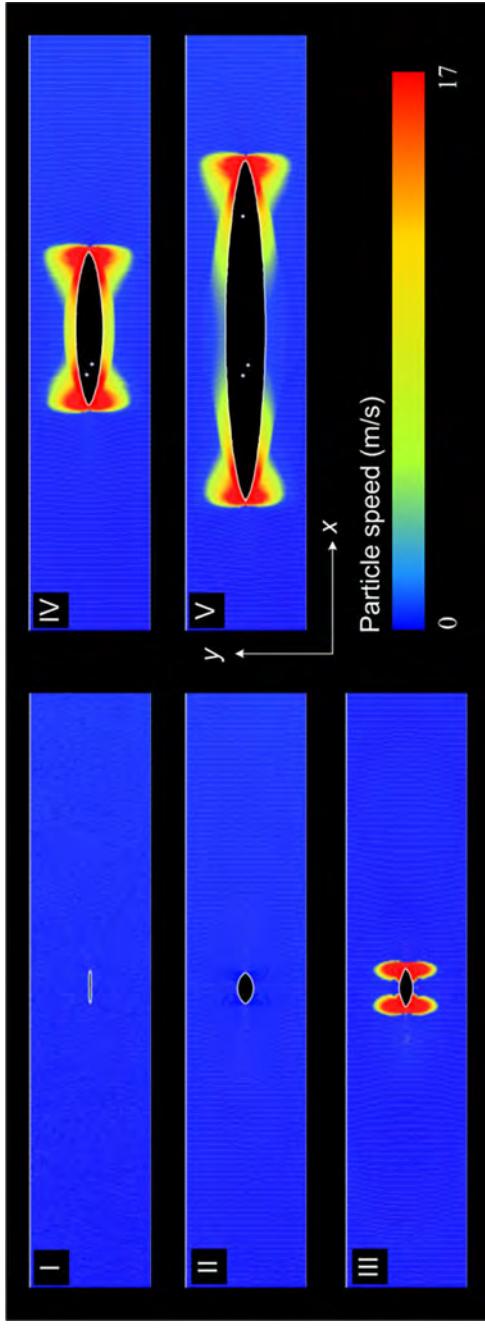
**Fig. 2** Snapshots of deformation for a meshwork under uniaxial tensile (mode I) loading. The meshwork is deformed in the  $y$ -direction, with zero strain in the  $x$ -direction ( $\varepsilon_x=0$ ) and I:  $\varepsilon_{y1}=37\%$ , II:  $\varepsilon_{y2}=161\%$ , III:  $\varepsilon_{y3}=162\%$ , IV:  $\varepsilon_{y4}=167\%$ , V:  $\varepsilon_{y5}=172\%$ , and VI:  $\varepsilon_{y6}=174\%$ ; constant strain rate of  $0.13$  ns $^{-1}$ . As loading is increased the crack shape changes from initially horizontally oriented (in I) to a vertically oriented shape (in II and following snapshots). As the crack propagates through the meshwork, two wave fronts emerge at the tip of the moving crack. One wave front is almost orthogonal to the crack propagation direction whereas the other one forms a well-defined cone. The angle  $\alpha$  denotes the angle between the wave front and the crack propagation direction, and resembles a Mach cone



**Fig. 3** Shear stress-strain relation at different axial loading  $\varepsilon_x$  (a) and shear modulus  $G$  as a function of axial loading  $\varepsilon_x$  (b). The results show that shear moduli  $G$  exponentially increases with increasing  $\varepsilon_x$ , and that a broad range of  $G$  can be achieved by altering  $\varepsilon_x$

26 m/s, and  $C=28$  m/s from small to large deformation, or from  $\varepsilon_{y3}$ ,  $\varepsilon_{y4}$ ,  $\varepsilon_{y5}$ , to  $\varepsilon_{y6}$ . We observe that as the crack moves through the material, two distinct wave fronts emanate from each crack tip as shown in Fig. 2(III–VI). One of the wave fronts resembles a shock front similar to a Mach cone, whereas the other one is almost orthogonal to the crack propagation direction. For the clearly visible shock wave, the measured angles between the wave front and the crack propagation direction are  $\alpha \approx 53^\circ$ ,  $45^\circ$ ,  $36^\circ$ , and  $33^\circ$  for strains  $\varepsilon_{y3}$ ,  $\varepsilon_{y4}$ ,  $\varepsilon_{y5}$ , and  $\varepsilon_{y6}$ , respectively. Using these angles we estimate the wave speed associated with this particular shock wave based on the relation  $\sin(\alpha) = C_0/C$ , where  $C_0$  is the characteristic wave speed that corresponds to the wave front with an angle of  $\alpha$ . Based on this analysis we find  $C_0 = 14$ , 15, 15, and 15 m/s of  $\varepsilon_{y3}$ ,  $\varepsilon_{y4}$ ,  $\varepsilon_{y5}$ , and  $\varepsilon_{y6}$ , respectively. The occurrence of the shock wave suggests that crack

propagation occurs faster than a characteristic wave speed of the material, which is estimated to be around 14–15 m/s. According to continuum theory the Rayleigh-wave speed is given by  $C_R \approx 0.923 \sqrt{G/\rho}$ , the shear-wave speed by  $C_S = \sqrt{G/\rho}$  and the longitudinal wave speed is given by  $C_L = \sqrt{3G/\rho}$ , where  $G$  is the shear modulus of material and  $\rho$  is the averaged mass density of the meshwork model. To estimate the wave speeds as a function of applied strain for the protein meshwork considered here we apply a shear stress  $\tau_{xy}$  on the boundary of a perfect meshwork model under tensile loading with a strain  $\varepsilon_x$  and record the shear stress over shear strain response as shown in Fig. 3a. We measure the shear modulus  $G$  of the meshwork material by computing the slope of the curve at the origin, with results shown in Fig. 3b. By continuously changing the value of  $\varepsilon_x$  we compute the shear modulus as a function of applied strain



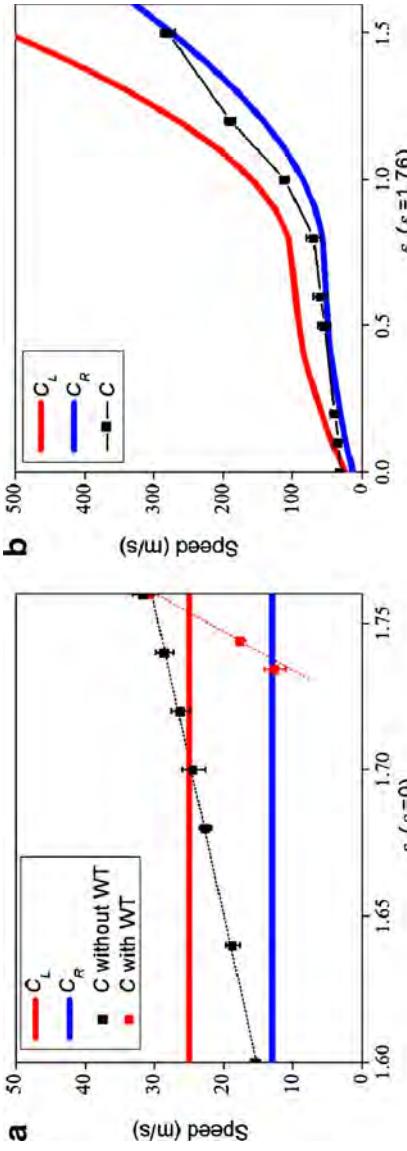
**Fig. 4** Snapshots of deformation for a meshwork under equi-biaxial tensile loading. The meshwork is loaded by equi-biaxial tension in the  $x$ - and  $y$ -directions (loading condition shown in Fig. 1b); constant strain rate of  $0.13 \text{ ns}^{-1}$ . The meshwork is deformed with an applied strains of *I*:  $\varepsilon_{xy1} = 37\%$ , *II*:  $\varepsilon_{xy2} = 162\%$ , *III*:  $\varepsilon_{xy3} = 161\%$ , *IV*:  $\varepsilon_{xy4} = 165\%$ , and *V*:  $\varepsilon_{xy5} = 168\%$ . The crack starts to propagate at the critical strain of  $\varepsilon_{xy3} = 162\%$ . Comparing to the uniaxial tension test shown in Fig. 2, there are two main differences: first, the initial  $x$ -direction crack does not flip to the  $y$ -direction before failure; second, no shock front or Mach cone is observed here

in the orthogonal direction. The analysis shows that the shear modulus and thus wave speeds increase as the loading is increased. For  $\varepsilon_x=0$  the shear-wave speed and longitudinal wave speed in the  $x$ -direction are estimated to be  $C_L=25$  m/s and  $C_S=14$  m/s, respectively. We thus conclude that the shock wave front seen here corresponds to the shear-wave speed that it is caused by crack propagation faster than the speed of shear waves in the surrounding meshwork material, where  $C>C_S$ , representing inter sonic crack propagation.

We suggest that the occurrence of inter sonic crack propagation is due to the stiffening of the meshwork in the loading direction ( $y$ -direction), which results in an increased tangent modulus and as a result, in increased wave speeds and thus faster energy transport that can drive the crack to faster velocities [24, 29]. The shock front emerges because the wave speeds become anisotropic under uniaxial loading. The shear-wave speed and thus rate of energy transport in the  $y$ -direction increases as loading is increased due to the stiffening of the meshwork. Yet, the shear-wave speed in the  $x$ -direction is limited to a constant value, and hence, we observe the emergence of a shock front as the crack speed is driven to larger values in excess of the shear-wave speed. To test this hypothesis we carry out an equi-biaxial fracture test where the applied strain is identical in the  $x$ - and the  $y$ -direction, such that  $\varepsilon_y=\varepsilon_x$ . We expect a lack of shock wave fronts in this case due to the fact that under equi-biaxial loading the tangent moduli and thus wave speeds increase equally in both the  $x$ - and  $y$ -directions. The dynamics of meshwork failure under equi-biaxial tensile loading is depicted in Fig. 4. Two differences are immediately noticeable by comparing to the uniaxial

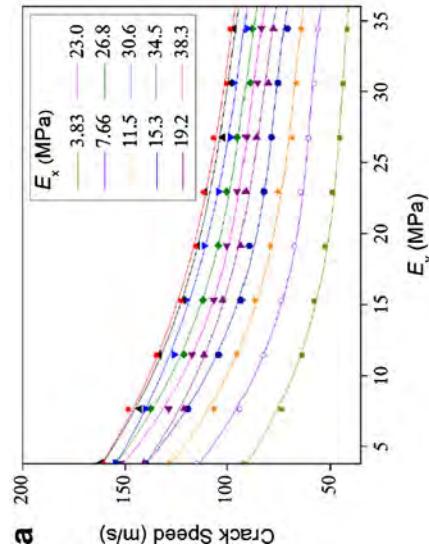
tension case shown in Fig. 2. First, the initial  $x$ -direction oriented crack does not flip to the  $y$ -direction. This is because under biaxial tension the strain in the filaments in both the  $x$ - and  $y$ -directions increase equally, preventing a change of the crack shape before failure. Second, in agreement with our hypothesis, no shock wave is observed upon failure initiation here in stark contrast to the case of uniaxial loading studied in Fig. 2.

The applied tensile strain is an important factor that affects the crack speed. For uniaxial loading ( $\varepsilon_x=0$ ), we observe that the crack speed  $C$  is higher than the shear-wave speed  $C_S$  and increases linearly with  $\varepsilon_y$  and can slightly exceed  $C_L$  at the ultimate strain beyond  $\varepsilon_y=172\%$  as shown in Fig. 5a. The maximum speed we observe under this loading condition is 32 m/s. We also investigate a case of multi-axial loading, where we keep  $\varepsilon_y=176\%$  and change  $\varepsilon_x$  (note this case is different from the equi-biaxial loading discussed above since here generally  $\varepsilon_y\neq\varepsilon_x$ ). We find that the crack speed increases significantly with increasing  $\varepsilon_x$  with a maximum speed of 280 m/s as shown in Fig. 5b. For comparison, we also plot the wave speeds  $C_R$  (the limiting speed predicted by linear-elastic fracture mechanics for a mode I crack) and  $C_L$  and find that they also increase with increasing strain in the  $x$ -direction, defining an upper and lower bound for the crack propagation speed. The observations made here show that the change in the applied strain in the nuclear lamina meshwork has significant implications on their dynamic fracture behavior. Indeed, it has been found in rupture experiments of rubber [17] that the crack propagation speed in a biaxial fracture test is between the  $C_R$  and  $C_L$  in



**Fig. 5** Steady-state crack propagation speeds, Rayleigh-wave speed and longitudinal wave speed as a function of loading conditions. **a** Crack speed as a function of uniaxial loading strain  $\varepsilon_y$ . The crack speed is calculated from the derivative of half crack length  $a$  to the time once the value becomes steady; the fluctuations of the measured speed define the standard deviation as shown by the error bars. The Rayleigh-wave speed and longitudinal wave speed in the  $x$ -direction are calculated from the shear modulus  $G$  (as given in Fig. 3b) of a perfect meshwork model subjected to the corresponding loading

condition. **b** Crack speed as a function of loading strain  $\varepsilon_y$ , with  $\varepsilon_x=176\%$  kept constant. The crack speed significantly increases as  $\varepsilon_x$  increases, and the crack speed is between the Rayleigh-wave speed and longitudinal wave speed in the plastic zone, which act as the lower and upper bound of the crack speed, respectively. The results shown here demonstrate that the crack propagation speed in the nuclear lamina is governed by the loading condition and the lateral strain  $\varepsilon_x$  is more dominant in governing the crack speed

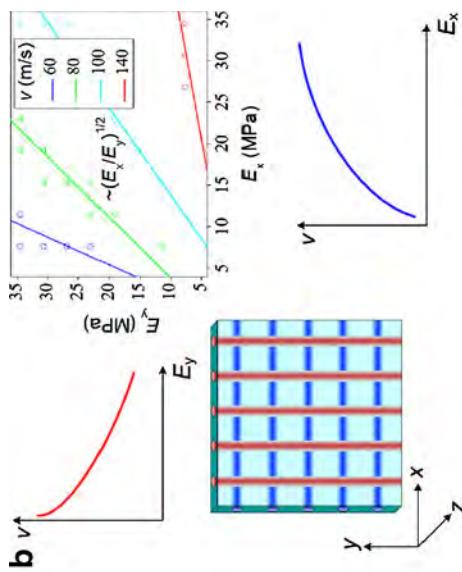


**Fig. 6** Crack speed in an ideal anisotropic elastic meshwork. As the schematic of the geometry of the meshwork and definition of coordinates shows in panel **b**, the meshwork is composed of two kinds of linear-elastic filaments with different stiffnesses in the two orthogonal directions, which results in an anisotropic material with different moduli in the  $x$ - and  $y$  directions ( $E_x$  and  $E_y$ , respectively). The ultimate failure strain of each filament is set to be 176%, and crack speeds in this study correspond to the steady-state obtained at applied strain of  $\varepsilon_x=0$  and  $\varepsilon_y=156\%$ . **a**, Crack propagation speed as a function of  $E_x$  and  $E_y$ . The maximum modulus (found to be 38.3 MPa) here corresponds to the maximum tangent modulus of the nuclear filament meshwork system

the rubber thin film, in agreement with our findings. For the specific cases considered here, the maximum speed reaches 22 times of the Rayleigh-wave speed at the relaxed zero-strain state.

We also study effect of damping due to the presence of water, by considering a viscous drag force. Under the presence of water the critical strain  $\varepsilon_C$  needed for crack initiation increases as the crack starts to propagate at a higher strain level. Once the crack starts to move, its steady speed starts with the Rayleigh-wave speed, increases in a steeper linear function with the increasing strain condition and reaches a maximum value as the same strain as in the model without solvent effect as shown in Fig. 5a. This result agrees with experimental observations that suggested that solvent viscosity slows down crack initiation as well as the propagation speed under constant pulling rates [30]. However, what we observed here is that the effect of viscosity vanishes at extreme loading conditions and does not affect the steady-state maximum crack speed in this case.

Unlike engineering materials such as metals, glass, and silicon, the nuclear lamina has a complex hierarchical structure as shown in Fig. 1. The change of the crack speed in the meshwork as a function of the applied strain is controlled by the stiffening behavior of individual intermediate filaments, which in turn is caused by characteristic changes in the molecular structure. As the secondary structure of each intermediate filament undergoes alpha-helix unfolding and an alpha-beta transition, the tangent



**Fig. 6** The crack speed decreases as  $E_y$  increases and increases as  $E_x$  increases, as illustrated by schematic plots in red and blue, respectively. Data points with same crack speeds are combined in groups as shown in the plot, which confirms that materials with same  $\sqrt{E_x/E_y}$  feature the same crack speed. The results demonstrate that the crack propagation speed is governed by the ratio of the tangent material modulus in and perpendicular to the crack propagation direction, as given by  $v \sim \sqrt{E_x/E_y}$ . Therefore various crack propagation speeds are accessible via different ratios of  $\sqrt{E_x/E_y}$  as can be achieved by different strain levels ( $\varepsilon_x$ ,  $\varepsilon_y$ ) in the real intermediate filament meshwork system

shear stiffness of the meshwork experiences a significant increase, leading to a greater  $C_R$ ,  $C_S$ , and  $C_L$  at larger deformation since  $C_{R,S,L} \sim \sqrt{G}$ . To directly test the hypothesis that the tangent stiffness controls crack dynamics, we measure the crack propagation speed in a linear-elastic meshwork system where we define a different stiffness for filaments in the  $x$ - and  $y$ -directions, respectively; where filaments are modeled by simple linear springs. This results in a meshwork with distinct elastic moduli  $E_x$  and  $E_y$  in the  $x$ - and  $y$ -directions, respectively. This mimics the anisotropic material properties that emerge when the loading in the  $x$ - and  $y$ -directions is different. Figure 6a shows the crack speed for varying combinations of elastic moduli, where crack propagation occurs in the  $x$ -direction and the moduli  $E_x$  and  $E_y$  are systematically changed. As can be inferred from Fig. 6a, the crack speed reaches the largest value as  $E_y$  decreases toward 0 MPa and  $E_x$  increases toward 38.3 MPa, which corresponds to the maximum tangent modulus of nuclear lamina before failure, as represented by Fig. 6b. The crack speed exponentially increases with the ratio  $\sqrt{E_x/E_y}$  for variations of the elastic moduli in the orthogonal directions. This suggests that the crack propagation speed is governed by the ratio of the tangent material moduli in and perpendicular to the crack propagation direction, as given by  $v \sim \sqrt{E_x/E_y}$ . Therefore, various crack propagation speeds are accessible via different ratios of  $\sqrt{E_x/E_y}$  as can be achieved by different strain levels ( $\varepsilon_x$ ,  $\varepsilon_y$ ) in the natural intermediate filament meshwork system.

## 4 Conclusion

Our results show that the speed of the crack propagation in intermediate filament meshwork is governed by the applied tensile strain in both the  $x$ - and  $y$ -directions. Under uniaxial loading tests, the crack speed exceeds the shear-wave speed, leading to a shock front or Mach cone at the crack tip, as shown in Fig. 2. We found that this phenomenon is caused by the material stiffening behavior, which is induced by the alpha-to-beta transition in the protein's secondary structure for the specific case considered here. This result is not limited to the particular material studied here but can have broad applications to many other protein or polymer materials with structural transitions and related stiffening (or softening) behavior under loading which may control the dynamics of material failure. Generally, the crack propagation speed reaches values greater than those predicted by simply considering the small-deformation moduli and associated wave speeds. It is noted that whereas intersonic crack speed contradicts conventional fracture theory for engineering materials, it agrees with the findings from earlier studies of nonlinear materials [24, 29, 31, 32].

The rapid propagation speed of the nuclear lamina meshwork seen here may reflect the behavior of cell nuclei during the cell division process from mechanical points of view. The occurrence of rapid fracture shows that once an extreme loading condition is reached, cracks propagate at a fast (here: intersonic) speed. This behavior may explain experimental observations that the nuclear envelope breakdown processes during cell divisions are indeed found to occur at short time-scales and are irreversible. For human somatic cells, the chromatin is tightly associated with the nuclear lamina [21, 22], which may by some extent slow down crack propagation speeds. However, neither *in situ* nor *in vitro* experiments of somatic nuclear laminas are available yet. We hope that our study will trigger a broader interest in studying dynamic failure of cell nuclei, which may eventually facilitate the development of more complex somatic nuclear lamina models through the availability of additional experimental data.

More generally, the multi-scale method used here provides a powerful bridge to overcome the four orders of magnitudes difference in length scales to appropriately describe the characteristic of catastrophic failure in the nuclear lamina meshwork. Based on this method, it may also be possible to study the effect of diseases caused by point mutations and their effect on the mechanical and failure properties of nuclear laminas. The approach used here could also be applied to other hierarchical materials and provide insight into their dynamic failure mechanisms [33–37]. Additional work is needed to develop a solid theoretical foundation for fracture in discrete nonlinear systems as studied here and to provide a

deeper investigation of the effect of solvent and chromatin on crack dynamics.

## References

- Herrmann, H., Bar, H., Kreplak, L., Strelkov, S. V., Aeby, U. (2007). Intermediate filaments: from cell architecture to nanomechanics. *Nature Reviews Molecular Cell Biology*, 8(7), 562–573.
- Aeby, U., Cohn, J., Buhle, L., Gerace, L. (1986). The nuclear lamina is a meshwork of intermediate-type filaments. *Nature*, 323(6088), 560–564.
- Ishikawa, H., Bischoff, R., Holtzer, H. (1968). Mitosis and intermediate-sized filaments in developing skeletal muscle. *The Journal of Cell Biology*, 38(3), 538–555.
- Chang, L., Shav-Tal, Y., Trcek, T., Singer, R. H., Goldman, R. D. (2006). Assembling an intermediate filament network by dynamic cotranslation. *The Journal of Cell Biology*, 172(5), 747–758.
- Omary, M. B., Coulombe, P. A., McLean, W. H. I. (2004). Mechanisms of disease: intermediate filament proteins and their associated diseases. *The New England Journal of Medicine*, 351(20), 2087–2100.
- van der Kooi, A. J., Bonne, G., Eymard, B., Duboc, D., Talim, B., Van der Valk, M., et al. (2002). Lamin A/C mutations with lipodystrophy, cardiac abnormalities, and muscular dystrophy. *Neurology*, 59(4), 620–623.
- Dahl, K. N., Scaffidi, P., Islam, M. F., Yodh, A. G., Wilson, K. L., Misteli, T. (2006). Distinct structural and mechanical properties of the nuclear lamina in Hutchinson-Gilford progeria syndrome. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 103(27), 10271–10276.
- Eriksson, M., Brown, W. T., Gordon, L. B., Glynn, M. W., Singer, J., Scott, L., et al. (2003). Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature*, 423(6937), 293–298.
- Djacek, W., Starzyk, H., Rzucidlo, Z. (1973). X-ray-irradiation induced changes of nuclear membrane of Kirkman-Robbins tumor-cells. *Experientia*, 29(1), 83–84.
- Strelkov, S. V., Schumacher, J., Burkhardt, P., Aeby, U., Herrmann, H. (2004). Crystal structure of the human lamin A coil 2B dimer: implications for the head-to-tail association of nuclear lamins. *Journal Of Molecular Biology*, 343(4), 1067–1080.
- Buehler, M. J., & Yung, Y. C. (2009). Deformation and failure of protein materials in physiologically extreme conditions and disease. *Nature Materials*, 8(3), 175–188.
- Scaffidi, P., & Misteli, T. (2006). Lamin A-dependent nuclear defects in human aging. *Science*, 312(5776), 1059–1063.
- Herrmann, H., & Aeby, U. (1998). Intermediate filament assembly: fibrilllogenesis is driven by decisive dimer-dimer interactions. *Current Opinion in Structural Biology*, 8(2), 177–185.
- Carpinteri, A., & Pugno, N. M. (2008). Mechanics of hierarchical materials. *International Journal of Fracture*, 150(1–2), 221–226.
- Ackbarow, T., Sen, D., Thaulow, C., Buehler, M. J. (2009). Alpha-helical protein networks are self-protective and flaw-tolerant. *PLoS ONE*, 4(6), e6015.
- Qin, Z., Buehler, M. J. (2011). Flaw tolerance of nuclear intermediate filament lamina under extreme mechanical deformation. *ACS Nano*. doi:[10.1021/mn200107u](https://doi.org/10.1021/mn200107u).
- Petersan, P. J., Deegan, R. D., Marder, M., Swinney, H. L. (2004). Cracks in rubber under tension exceed the shear wave speed. *Physical Review Letters*, 93(1), 015504.
- Sen, D., Thaulow, C., Schieffer, S. V., Cohen, A., Buehler, M. J. (2010). Atomistic study of crack-tip cleavage to dislocation emission in discrete nonlinear systems as studied here and to provide a

transition in silicon single crystals. *Physical Review Letters*, 104(23), 235502.

19. Qin, Z., Kreplak, L., Buehler, M. J. (2009). Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. *PLoS ONE*, 4(10), e7294.
20. Qin, Z., & Buehler, M. J. (2010). Molecular dynamics simulation of the alpha-helix to beta-sheet transition in coiled protein filaments: evidence for a critical filament length scale. *Physical Review Letters*, 104(19), 198304.
21. Goldberg, M. W., Hattenbach, I., Hutchison, C. J., Stick, R. (2008). Filaments made from A- and B-type lamins differ in structure and organization. *Journal of Cell Science*, 121(2), 215–225.
22. Dahl, K. N., Kahn, S. M., Wilson, K. L., Discher, D. E. (2004). The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. *Journal of Cell Science*, 177(20), 4779–4786.
23. Bertaude, J., Qin, Z., Buehler, M. J. (2010). Intermediate filament-deficient cells are mechanically softer at large deformation: a multi-scale simulation study. *Acta Biomaterialia*, 6(7), 2457–2466.
24. Buehler, M. J., & Gao, H. J. (2006). Dynamical fracture instabilities due to local hyperelasticity at crack tips. *Nature*, 439(7074), 307–310.
25. Batchelor, G. K. (1967). *An introduction to fluid dynamics*. Cambridge: University Press.
26. Jennings, B. R., & Parslow, K. (1988). Particle-size measurement—the equivalent spherical diameter. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 419(1856), 137–149.

27. Plimpton, S. (1995). Fast parallel algorithms for short-range molecular-dynamics. *Journal of Computational Physics*, 117(1), 1–19.
28. Humphrey, W., Dalke, A., Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33–38.
29. Buehler, M. J., Abraham, F. F., Gao, H. J. (2003). Hyperelasticity governs dynamic fracture at a critical length scale. *Nature*, 426(6963), 141–146.
30. Baumberger, T., Caroli, C., Martina, D. (2006). Solvent control of crack dynamics in a reversible hydrogel. *Nature Materials*, 5(7), 552–555.
31. Pugno, N. M. (2006). Dynamic quantized fracture mechanics. *International Journal of Fracture*, 140(1–4), 159–168.
32. Pugno, N. M., & Ruoff, R. S. (2004). Quantized fracture mechanics. *Philosophical Magazine*, 84(27), 2829–2845.
33. Wegst, U. G. K., & Ashby, M. F. (2004). The mechanical efficiency of natural materials. *Philosophical Magazine*, 84(21), 2167–2181.
34. Espinosa, H. D., Rim, J. E., Barthelat, F., Buehler, M. J. (2009). Mergers of structure and material in nacre and bone—perspectives on de novo biomimetic materials. *Progress in Materials Science*, 54(8), 1059–1100.
35. Nukala, P. K., & Simunovic, S. (2005). Statistical physics models for nacre fracture simulation. *Phys Rev E Stat Nonlin Soft Matter Phys*, 72(4 Pt 1), 041919.
36. Pugno, N. M. (2006). Mimicking nacre with super-nanotubes for producing optimized super-composites. *Nanotechnology*, 17(21), 5480–5484.
37. Qin, Z., Feng, X. Q., Zou, J., Yin, Y. J., Yu, S. W. (2007). Superior flexibility of super carbon nanotubes: molecular dynamics simulations. *Applied Physics Letters*, 91(4), 043108.

# Hierarchical Structure Controls Nanomechanical Properties of Vimentin Intermediate Filaments

Zhao Qin<sup>1,2</sup>, Laurent Kreplak<sup>3</sup>, Markus J. Buehler<sup>1,2,4\*</sup>

**1** Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **2** Center for Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **3** Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada, **4** Center for Computational Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America

## Abstract

Intermediate filaments (IFs), in addition to microtubules and microfilaments, are one of the three major components of the cytoskeleton in eukaryotic cells, playing a vital role in mechanotransduction and in providing mechanical stability to cells. Despite the importance of IF mechanics for cell biology and cell mechanics, the structural basis for their mechanical properties remains unknown. Specifically, our understanding of fundamental filament properties, such as the basis for their great extensibility, stiffening properties, and their exceptional mechanical resilience remains limited. This has prevented us from answering fundamental structure-function relationship questions related to the biomechanical role of intermediate filaments, which is crucial to link structure and function in the protein material's biological context. Here we utilize an atomistic-level model of the human vimentin dimer and tetramer to study their response to mechanical tensile stress, and describe a detailed analysis of the mechanical properties and associated deformation mechanisms. We observe a transition from alpha-helices to beta-sheets with subsequent interdimer sliding under mechanical deformation, which has been inferred previously from experimental results. By upscaling our results we report, for the first time, a quantitative comparison to experimental results of IF nanomechanics, showing good agreement. Through the identification of links between structures and deformation mechanisms at distinct hierarchical levels, we show that the multi-scale structure of IFs is crucial for their characteristic mechanical properties, in particular their ability to undergo severe deformation of  $\approx 300\%$  strain without breaking, facilitated by a cascaded activation of a distinct deformation mechanisms operating at different levels. This process enables IFs to combine disparate properties such as mechanoinsensitivity, strength and deformability. Our results enable a new paradigm in studying biological and mechanical properties of IFs from an atomistic perspective, and lay the foundation to understanding how properties of individual protein molecules can have profound effects at larger length-scales.

**Citation:** Qin Z, Kreplak L, Buehler MJ (2009) Hierarchical Structure Controls Nanomechanical Properties of Vimentin Intermediate Filaments. PLoS ONE 4(10): e7294. doi:10.1371/journal.pone.0007294

**Editor:** Joel M. Schnur, George Mason University, United States of America

**Received:** June 25, 2009; **Accepted:** August 21, 2009; **Published:** October 6, 2009

**Copyright:** © 2009 Qin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** ZQ and MJB acknowledge support by the Air Force Office of Scientific Research (grant # FA9550-08-1-0321) and the National Science Foundation (grant # DMR-081976). This research was supported by an allocation of advanced computing resources supported by the National Science Foundation (Teragrid, grant # MSS090007). The authors further acknowledge support from the TeraGrid Advanced Support Program. LK acknowledges support by the National Sciences and Engineering Research Council of Canada. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

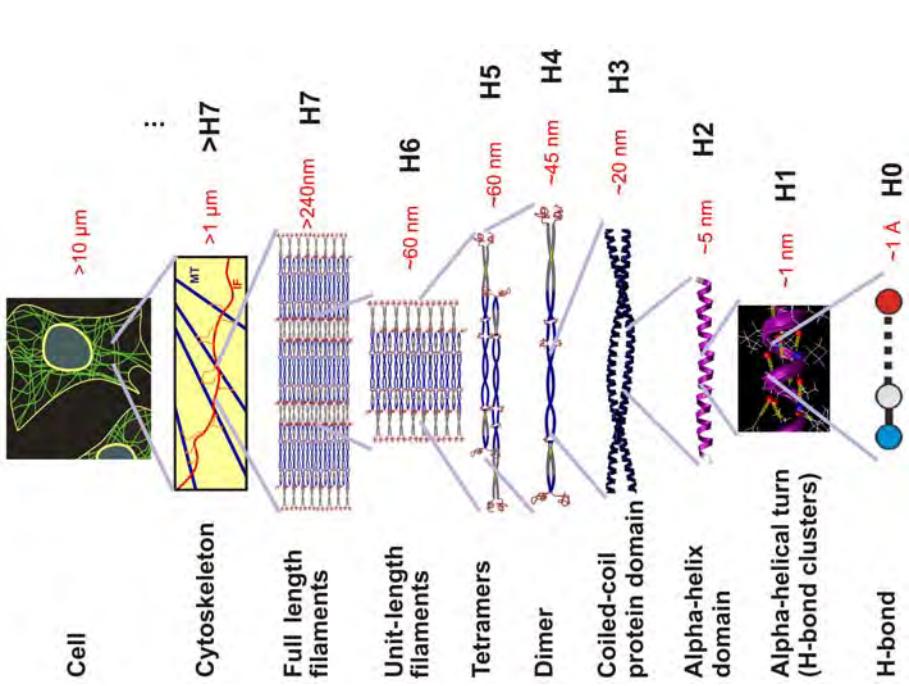
\* E-mail: mbuehler@mit.edu

## Introduction

Intermediate filaments (IFs), in addition to microtubules and microfilaments, are one of the three major components of the cytoskeleton in eukaryotic cells [1,2,3]. IFs are crucial in defining key biomechanical functions of cells such as cell migration, cell division and mechanotransduction, and have also been referred to as the “safety belts of cells” to prevent exceedingly large cell stretch [4,5,6]. Earlier studies focused on analyzing the mechanical signature of IFs have suggested that they are highly sensitive to applied forces at small levels, and that they can sustain extremely large deformation of up to 300% [3,7,8]. It has also been observed that due to severe stiffening, the tangent modulus of IFs increases manifold during deformation, a property that is believed to be crucial for providing mechanical stability to cells at large stretch. IFs are also found in the cell's nucleus in the form of laminin, where they form a dense mesh-like network providing mechanical

integrity and biochemical functions at the cytoskeleton-chromatin interface [9,10,11]. IFs have also been associated with dozens of genetic diseases, where single point mutations and deletions lead to structural changes at different levels in the IFs organization. If related diseases include muscle dystrophies, Alexander disease [12], epidermolysis bullosa simplex, as well as a broad class of disorders referred to as laminopathies (e.g. rapid aging disease progeria) [13].

IFs form hierarchical structures as shown in Fig. 1, ranging from H-bonds, clusters of H-bonds in alpha-helical turns, alpha-helical proteins, dimers, tetramers, unit length filaments, full-length filaments to the cellular level. Each vimentin dimer contains 466 amino acids and forms the fundamental building block of IFs. Each dimer consists of four major structural segments linked in series in the sequence 1A, 1B, 2A and 2B, connected by linkers L1, L12 and L2 (Fig. 2A). Although there exists clear evidence that IFs play a key role in important physiological and pathological



**Figure 1. The hierarchical structure of intermediate filaments, from atomic to cellular scales.** The figure shows relevant structural levels H0–H7. See Table 1 for deformation mechanisms associated with each hierarchical level.  
doi:10.1371/journal.pone.0007294.g001

mechanisms, a complete atomistic-level molecular model of the basic constituents of this kind of protein material remains elusive. So far, only parts of two of the four segments, a section of the 1A and 2B domain of the vimentin dimer structure, have been crystallized and their atomic structure identified based on x-ray diffraction experiments [14,15,16,17] (the structures are found in Protein Data Bank (PDB) entries 1gk4, 1gk6 and 1gk7). There are persistent experimental challenges in identifying the remaining parts of the vimentin IF structure. Since IFs are intrinsically disordered structures [18], x-ray diffraction studies on naturally occurring or recombinantly produced IF bundles do not provide sufficient data to produce a full atomistic model of IFs. Solid state nuclear magnetic resonance has been successfully utilized to derive atomistic models of amyloid fibrils [19]. However, amyloid peptides are much smaller than IF dimers, so that this approach may remain unsuccessful as well. Finally, cryo-electron tomography offers the promise of molecular-level imaging of single IFs, but the best tomograms are currently limited to a resolution around 5 nm [20].

Due to the lack of a structural model of IFs, little is known of the biological and physical material concepts of intermediate filaments. Specifically, our understanding of fundamental filament properties, such as the basis for their great extensibility, stiffening properties, and their exceptional mechanical resilience remains limited. This has prevented us from answering fundamental

structure-function relationship questions related to the mechanical role of intermediate filaments, which is crucial to link structure and function in the protein material's biological context. The availability of a structural model of IFs is also the key to understand the mechanisms of IF related genetic diseases, where structural flaws cause major structural changes of biologically relevant properties. Here we utilize an atomistic-level model of the human vimentin dimer and tetramer to pursue a bottom-up nanomechanical analysis of the structure. In order to identify their nanomechanical properties at different levels in the material's hierarchical organization [21], we carry out a series of systematic molecular dynamics simulations to probe the response of the vimentin IF to mechanical tensile stretch. Specifically, the aim of our investigations is to elucidate the basis for their great extensibility, stiffening properties, and their exceptional mechanical resilience as well as associated molecular deformation mechanisms.

## Results

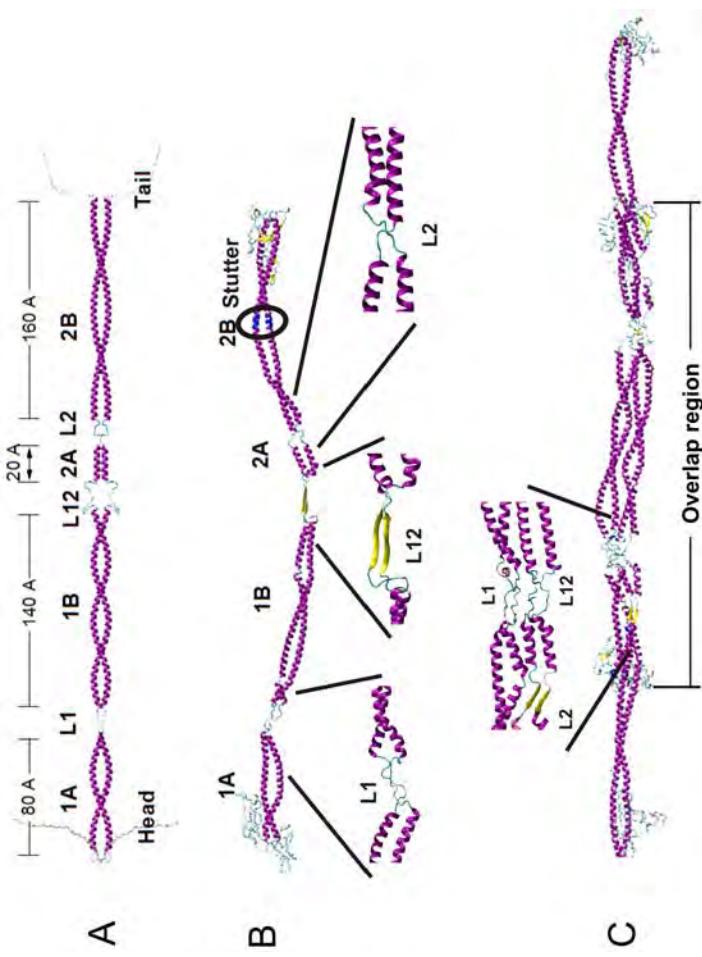
We apply a hierarchy of *in silico* techniques with different levels of accuracy, ranging from atomistic to coarse-grained representations, in order to enable a bottom-up multi-scale analysis of vimentin IFs at different hierarchical levels and time-scales. Atomistic level molecular dynamics simulations are carried out using the CHARMM19 all-atom energy function with an effective Gaussian model for the water solvent [22,23], as well as by using explicit solvent simulations with the CHARMM force field and TIP3 water [24] implemented in NAMD [25] for validation of the effective Gaussian simulations. Additional validation simulations are carried out using the MARTINI force field [26], used to develop a coarse-grained representation of the proteins. Details are included in the Materials and Methods section.

## Molecular structure of vimentin dimer and tetramer

We first present the geometries and associated visualizations of the equilibrated geometries of both the dimer and the tetramer vimentin IF. We then continue with a detailed structural analysis of the resulting structures and compare with experimental measurements.

Figure 2B shows a snapshot of the molecular model of the vimentin dimer, displaying the characteristic segmented geometry with coiled-coil regions connected through linker domains. The total length of the dimer without the head and tail domain is ≈48.7 nm. This result is in agreement with experimental results where a range of 46–49 nm has been reported earlier [15,27]. From dynamical analyses of the molecular dynamics trajectory of the dimer at 300 K, we observe that the linker domains represent soft, hinge-like structures that connect much stiffer coiled-coil segments (see Movie S1). This heterogeneous distribution of the bending stiffness along the molecule's axis strongly affects the filaments' motion by enhancing its mobility due an effectively reduced persistence length, which might be significant for IF self-assembly. Thereby, the linker domains act as hinges around which the molecular structure can rather easily bend and rotate.

Figure 2C shows a snapshot of the molecular model of the vimentin tetramer, consisting of two anti-parallel dimers [15]. The total length of the tetramer is ≈61.3 nm, which agrees with the length of unit length filaments observed in experiment (62 nm at pH 7.5 [27]). The overlap part of the tetramer has a length of ≈36 nm, where the experimental value is 30–36 nm [15,27]. Segments 1A, 1B and 2A are fully contained in this part, but the 2B segments in the immediate vicinity of the two terminals are located completely outside. We find that the head segment of each



**Figure 2. Molecular structures of vimentin dimer and tetramer (schematic and results of molecular simulation).** Panel A: Schematic geometry of the dimer geometry, including labels identifying the various segments and linker domains. Panel B: Snapshot of the atomistic model of the dimer. The blue part denotes the stutter part around the stutter (amino acids 345–351). Panel C: Snapshot of the atomistic model of the tetramer. The blow-ups in panels B and C show detailed views of the molecular structure in the dimer and tetramer. doi:10.1371/journal.pone.0007294.g002

dimer is coiled around the other dimer, increasing the contact surface area and thereby providing enhanced interdimer interactions. This structure is an alternative configuration to the sandwich model that has been discussed as a possible tetramer arrangement, where the head domain is located in between the two dimers [28]. Movie S2 shows the equilibrated structure of the vimentin tetramer, suggesting that the overlapped L1/L12 linker domains represent regions of locally soft bending stiffness similar to the dimer geometry. From an analysis of the movie it can be seen that compared with the dimer structure, the tetramer structure is notably more stiff due to its increased area moment of inertia, resulting in an effectively increased persistence length. Atomistic-level geometry PDB files of the dimer and tetramer can be found in Structure S1 (dimer) and Structure S2 (tetramer).

In both protein structures, the 1A segment contains 61 amino acids (residues 78–138), for which both phenylalanine and tyrosine residues are packed consecutively at the hydrophobic core positions near the segment center [29]. Therefore, the 1A segment is only marginally stable, which increases the mobility of the head domain during IF assembly as suggested in [17]. Segment 1B (residues 147–247) contains 101 amino acids with abundant hydrophobic residues distributed periodically around each alpha-helix. Therefore, the two chains generate a stable coiled-coil geometry, in agreement with results based on analytical ultracentrifugation [14]. Segment 2A (residues 264–282) is the shortest and the least flexible among all segments. It assumes the configuration of a nearly parallel alpha-helical bundle, in agreement with earlier suggestions [14]. Segment 2B with 115 amino acids (residues 291–405) features an alpha-helical coiled-coil geometry for the major segment, and a stutter region in the vicinity of residue 349. In the stutter region, the two alpha-helical chains unwind and give rise to a much longer coiled-coil pitch, resembling two parallel alpha-

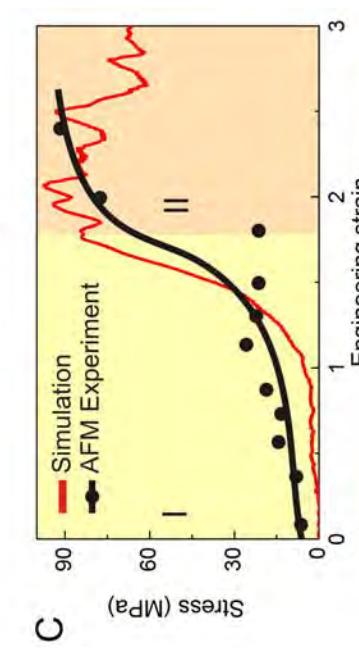
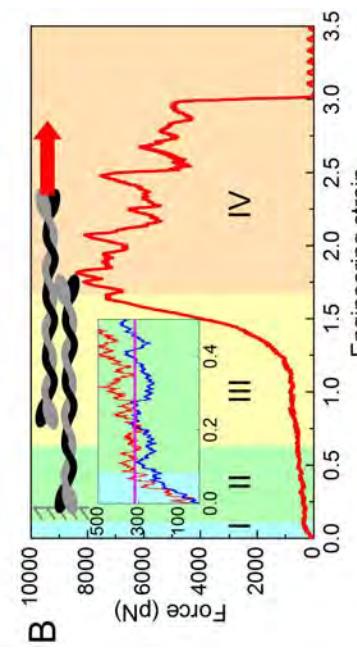
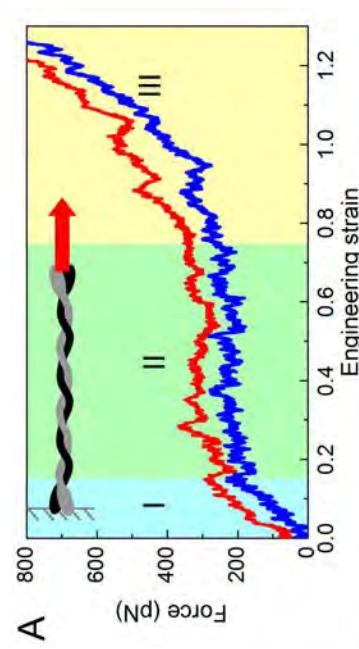
helices in that region (see Fig. 2B, region marked in blue). This regions is part of a segment that has already been crystallized, which also shows the formation of the stutter region [16,30], confirming our predictions. The left end of the 1A segment connects to the head domain, while the right end of the 2B segment connects to the tail domain [31], where these two flexible domains are found to be poorly structured, in agreement with earlier suggestions [14,17]. Linker L1 of 8 amino acids in length is among the softest regions in the structure, and facilitates the anti-parallel assembly of the dimers into larger-scale filaments, where these linkers align with the L12 domain in a regular pattern [32]. Linker L12 forms a beta-sheet structure in the dimer. An earlier sequence analysis has pointed out that the L12 domain region should form a small beta-sheet of two chains, with one face of the beta-sheet being apolar [32]. Our simulation confirms this hypothesis. Linker L2, which always features 8 amino acids in length for all IFs, results in a transition between two rod-like segments [32].

The average diameter of the cross section of the dimer is  $d_{\text{dimer}} \approx 2$  nm, and the average diameter of the tetramer is measured to be  $d_{\text{tetramer}} \approx 3.6$  nm, in close agreement with experimental measurements of 3.4 nm [15] (the small difference may be caused by the interaction with other dimers in larger filaments that contain many dimers, which may lead to increased compaction of the structure). Based on these geometric measurements we estimate that the diameter of a non-compacted unit length filament is  $d_{\text{ULF}} \approx 18.3$  nm, which is closely within the range measured in experiment, where values of 16–20 nm have been reported [33]. For compacted vimentin IFs we predict a diameter of  $d_{\text{IF,mature}} \approx 11.2$  nm, which is in excellent agreement with various experimental reports of values in the range of 10–12 nm [33,34]. A summary of the structural comparison between

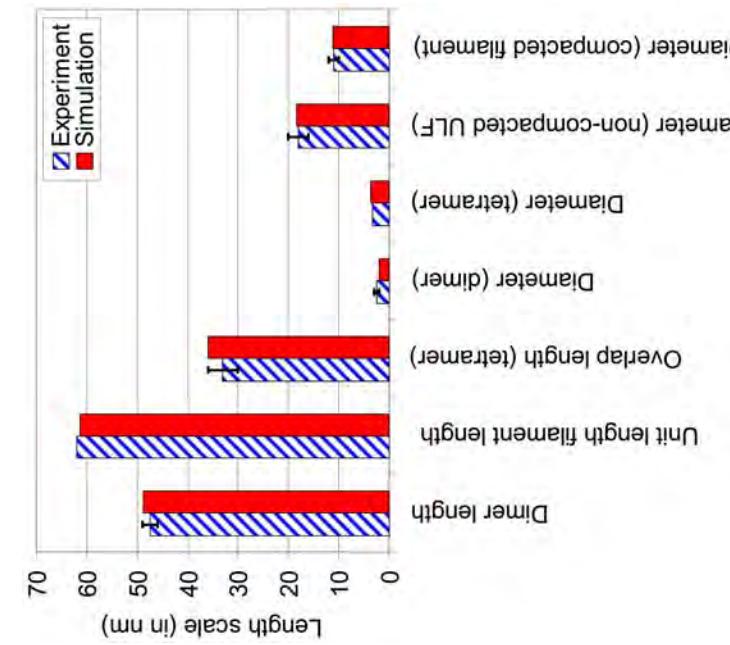
simulation predictions and experimental results is shown in Figure 3 (and Table S1).

**Nanomechanical analysis of vimentin dimer and tetramer**  
 We now proceed with a study of the nanomechanical properties of the dimer and tetramer structure by carrying out an engineering tensile test (for the pulling geometry, see inlays of Figs. 4A/B). Both structures are stretched until all helical domains are fully extended and/or failure due to intermolecular shear occurs. We begin with an analysis of the mechanical behavior of the dimer. Figure 4A shows two characteristic force-strain curves for two pulling speeds (0.01 Å/ps and 0.1 Å/ps). The simulations reveal three distinct regimes of deformation. In the first regime (I), the pulling force increases linearly with strain until it reaches an angular point, where a dramatic change in the slope emerges. In the second regime (II), the force-strain curve resembles a plateau regime where the pulling force remains almost constant. It is found that unfolding of all alpha-helical domains occurs in this regime (II). In the third regime (III), the stretching force increases rapidly as the strain increases. This is caused by pulling the unfolded polypeptide backbone of the protein, where the stretching of covalent bonds leads to a much increased stiffness. The overall characteristics of the force-strain curves of the dimer share similarities with stretching of individual alpha-helices [35,36] as well as with myosin coiled-coil structures [37], which agrees with the fact that most parts of the dimer are composed of coiled-coils (see Fig. 2B). Regime (II) represents the main unfolding regime, where the average force in this regime approximates the protein's unfolding force.

We continue with an analysis of the tetramer mechanical properties. Figure 4B shows the force-strain curves of the tetramer. The simulations reveal four distinct regimes. In the first regime (I),



**Figure 4. Force-strain relations for tensile deformation of the vimentin dimer, tetramer, as well as stress-strain relationship of full-length filaments (including comparison with AFM nanomechanics experiments).** Panel A: Force-strain curve of the dimer (0.01 Å/ps and 0.1 Å/ps pulling speeds). The force-strain curve features three regimes: (I), The force increases linearly with strain until the rupture of the H-bonds and uncoiling of alpha-helices sets in. (II), A plateau of a constant force is found during which the unfolding process takes place. (III), The filament stiffens significantly due to stretching of the protein backbone. Panel B: Force-strain curve of the tetramer (0.1 Å/ps pulling speed; results for 0.01 Å/ps pulling speed are shown in the inlay for strains less than 50%). The pink bar in the inlay denotes the unfolding force for a dimer. The force-strain curve features four regimes: (I), The force increases linearly with strain until the rupture of the H-bonds and uncoiling of alpha-helices begins. (II), Plateau with approximately constant force, where the unfolding of 2B segment takes place. (III), Unfolding of the 1A, 1B and 2B segments. (IV), Interdimer sliding. Panel C: Stress-strain behavior of IFs, quantitative comparison between simulation and experiment. The black curve depicts experimental results of stretching single IFs (data taken from reference [8]). The two regimes (I) and (II) indicated by different shades refer to unfolding of alpha-helices and interdimer sliding between dimers beyond 175% strain, respectively.  
 doi:10.1371/journal.pone.0007294.g004



**Figure 3. Comparison of structural properties of the IF dimer and tetramer between experiment and simulation (see also Table S1).**  
 doi:10.1371/journal.pone.0007294.g003

the pulling force increases linearly with the strain until it reaches the angular point. In the second regime (II), the force-strain curve forms a plateau, where the pulling force keeps almost constant, which is similar to the case of the deformation properties of the dimer. From simulation snapshots, we find that the 2B segments unfold first, starting from the stiffer location, then expanding throughout the entire segment. In the third regime (III), we observe a significant transition in the force-strain behavior. During this transition, the pulling force changes from the low-force plateau ( $<1$  nN) to a peak value of 7.3 nN, marking the onset of the final regime (IV). Beyond the peak value, the force level in regime (IV) begins to decrease. The force plateau observed in Figs. 4A and B relates to the stepwise unfolding of the coiled-coil segments. We observe that at the end of the plateau regime, the dimer and tetramer behave differently. Once the dimer is completely uncoiled, the pulling force abruptly increases due to the stretching of the protein backbone. Whereas for the tetramer, the overlapped part in center begins to uncoil, and the pulling force approaches a much broader transition region, where the structure becomes significantly stiffer and eventual fails by interdimer sliding. The two dimers are completely detached at  $\approx 300\%$  strain (Fig. 4B). The comparison between the dimer and tetramer reveals a significant difference in the mechanical behavior.

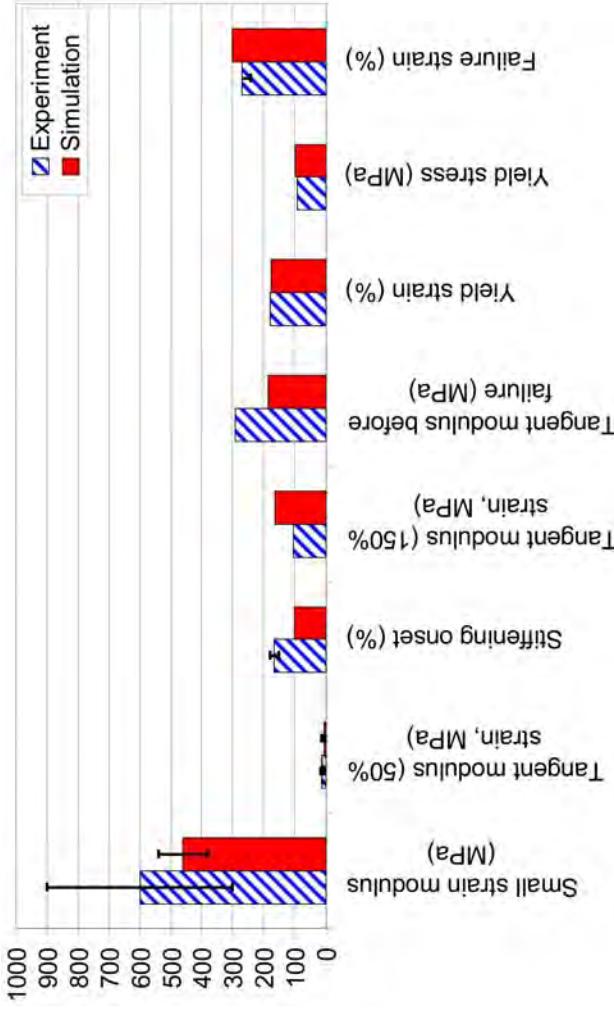
### Comparison with experiment

We present a quantitative comparison with AFM experiments of single IF filaments [8]. The comparison between simulation and experiment is shown in Fig. 4C, showing overall good agreement between simulation and experiment. Specifically, the Young's modulus for small deformation, calculated from the pre-angular point regime shown in Fig. 4A (where the dimer is deformed homogeneously), results in values in the range of  $\approx 380$ –540 MPa. This is in close agreement with experimental results from corresponding vimentin IF bending experiments, which show moduli in the range of 300–900 MPa [38]. This agreement supports the notion that IFs should be considered as a continuous body in the very small-deformation regime where no unfolding or

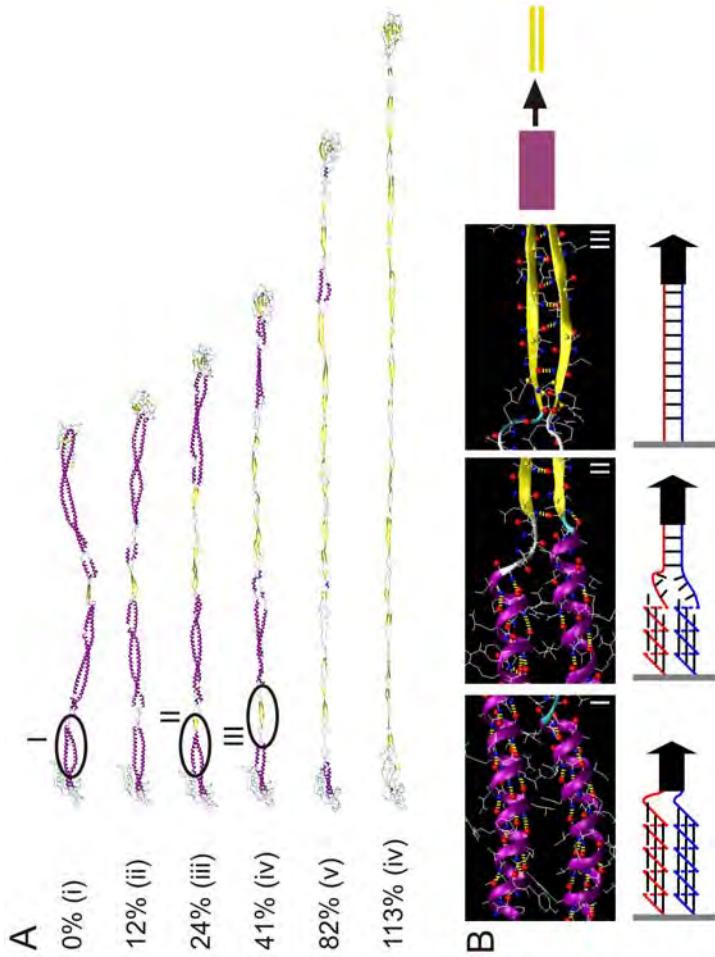
sliding occurs. The resulting maximum stress for the filament is predicted to be  $\approx 97$  MPa (found at  $\approx 205\%$  strain) from simulation (Fig. 4C), close to experimental measurements of 90 MPa at similar levels of strain [8]. The transitions of the stress-strain behavior and the maximum strength in the two methods show overall good agreement. However, in the lower strain regime before transition, the experimental stress is higher, possibly caused by effect of sliding of the filament on the surface during stretching. From the analysis of the stress-strain behavior, we extract the tangent modulus for the full length filament to be  $\approx 3$ –12 MPa below 50% strain. This modulus is close to experimental results, where a range of  $\approx 7$ –15 MPa was reported [8,34]. We note that the initial tangent Young's moduli found here are significantly smaller than typical values reported for  $\alpha$ -keratins (1.8 GPa), microtubules (0.9 GPa) [39] and F-actin (2.6 GPa) [40], confirming the characteristic properties of IFs. In both simulation and experiment, the tangent modulus increases by a factor of  $\approx 20$  once 150% strain is reached, leading to severe stiffening. The tangent modulus in simulation approaches  $\approx 185$  MPa before failure, similar to experimental results. Possible explanations for the discrepancies between experiment and simulation include the presence of defects, structural irregularities, and the fact that much longer filaments are considered in experiment, whereas the small system considered in our simulations is perfect without any structural flaws. A summary of the nanomechanical properties and a comparison between simulation and experiment is shown in Figure 5 (see also Table S2).

### Multi-scale deformation and failure mechanisms

Our simulations of mechanical deformation across several orders of magnitudes of strain until failure occurs enable us to carry out a detailed mechanistic and structural analysis of deformation mechanisms of both the dimer and tetramer. We begin by analyzing cartoon visualizations of the filament structures at different levels of applied strain. Figure 6A displays the structural deformation during the stretching simulation of the dimer. Once tensile load is applied, the four segments align in the pulling direction, and the coiled-coil protein domains are stretched slightly to mediate the elongation. No rupture or domain



**Figure 5. Comparison of mechanical properties of IF filaments between experiment and simulation (based on data shown in Table S2).**  
doi:10.1371/journal.pone.0007294.g005



**Figure 6. Simulation snapshots and structural analysis during pulling of a dimer.** Panel A displays snapshots as the applied tensile strain is increased up to 113%. Panel B depicts snapshots of the right part of the 1A segment as the  $\alpha$ - $\beta$  transition occurs (frames I to III) correspond to those highlighted in panel A). The schematics in the lower part describe the atomistic details of the  $\alpha$ - $\beta$  transition mechanism.

unfolding is observed until the applied strain is larger than 12%, where the 2A segment begins to uncoil first. The uncoiling of alpha-helical structures occurs simultaneously with the formation of beta-sheets, a phenomenon referred to as  $\alpha$ - $\beta$  transition. The pulling force causes the two uncoiled chains to come sufficiently close to form a beta-sheet protein domain. Figure 6B shows atomistic-level details of the  $\alpha$ - $\beta$  transition process, illustrating the mechanisms by which beta-sheets are formed due to uncoiling of pairs of alpha-helical domains (results shown for the right part of the 1A segment). The  $\alpha$ - $\beta$  transition as observed here has also been observed in experiment, where it was shown that the secondary structure of double-stranded coiled coil proteins gradually transform into beta-sheets under applied tensile load [34,41,42].

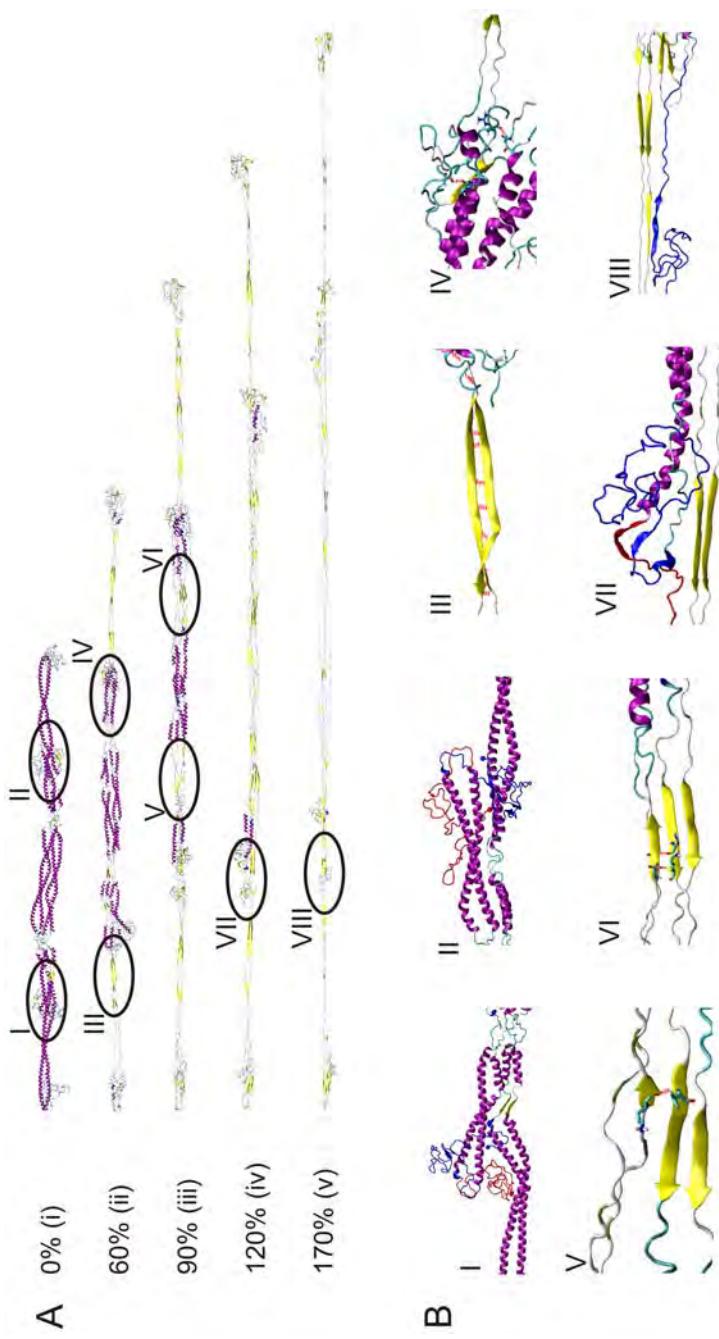
Figure 7A shows the structural characteristics of the tetramer under tensile deformation, showing molecular structure snapshots as the tetramer is being stretched. As the applied strain increases to 20%, the first part to unfold is the free-standing 2B segment (the  $\alpha$ - $\beta$  transition is also observed here). We find that at low strains there is no unfolding in the overlapped part, because the head part of each dimer forms H-bonds with the other dimer. Therefore, the two dimers are strongly connected in the head-coiled-coil overlap region. Once strain is increased to more than 60%, unfolding of the 1A segment in each of the two dimers begins, accompanied by beta-sheet formation. The anti-parallel beta-sheets formed by individual coiled-coil uncoiling processes interact with each other to form larger beta-sheet crystals as shown in Figs. 7B (see blow-ups). The three segments 1A, 1B, 2A begin to unfold sequentially between  $\approx$ 60% and  $\approx$ 100% strain. Unraveling of the 1A segment, the first to rupture at  $\approx$ 60% strain, stops at  $\approx$ 170% strain, requiring the largest applied strain among all three domains.

Further structural analysis is shown in Fig. 8, where we plot the number of amino acids associated with alpha-helical and beta-

sheet secondary structures as a function of the applied tensile strain in a geometry-strain map. We first focus on the dimer structure. The results shown in Fig. 8A confirm that the number of amino acids in the alpha-helical state is approximately constant for strains smaller than 12%, which demonstrates that the structure is intact.

The reason why the 2A segment ruptures first is that while the other three segments are coiled-coils, the 2A segment is essentially a parallel alpha-helical bundle, which has a lower mechanical stability against tensile stretch as pointed out earlier [14,21]. Once the 2A segment is fully uncoiled at  $\approx$ 19% strain, the other segments 1B, 1A and 2B begin to uncoil sequentially. Furthermore, the geometry-strain map shown in Fig. 8A (upper part) confirms that the segments do not unfold simultaneously, but rather in a sequential process. The observed sequence with unfolding in the order 2A  $\rightarrow$  1B  $\rightarrow$  1A  $\rightarrow$  2B is repeatable in different simulations with different pulling rates. As depicted in Fig. 8A, while the number of amino acids associated with alpha-helices decreases from  $\approx$ 520 to 0 between 12% and 110% strain, the number of amino acids associated with beta-sheets increases from  $\approx$ 10 to 370, supporting the notion that most amino acids in the dimer under pulling turn from alpha-helical dominant structures to beta-sheet structures (the  $\approx$ 10 beta-sheets associated amino acids in the beginning correspond to the short beta-strand found in the L12 linker).

The behavior of the tetramer is more complex. As shown in Fig. 8B, here the 2A domain unfolds fastest, followed by the 1B domain. The unfolding sequence with unfolding in the order 2B  $\rightarrow$  1A  $\rightarrow$  1B  $\rightarrow$  2A is repeatable in different simulations. Notably, the unfolding sequence in the tetramer is reversed when comparing it with the dimer. Once the 1A segment has completed the  $\alpha$ - $\beta$  transition, the structure is invariant beyond this region, as shown in Fig. 8B. It is also noted that unlike as in the case of



**Figure 7. Simulation snapshots and structural analysis during pulling of the tetramer.** Panel A displays snapshots as the applied tensile strain is increased up to 170%. The  $\alpha/\beta$  transition leads to significant thinning of the filament diameter (compare snapshot (i) with snapshot (v)). Panel B displays blow-ups of detailed atomic structures of the snapshots marked (i) to (vii) in panel A.  
doi:10.1371/journal.pone.0007294.g007

stretching the dimer, the sum of the number of amino acids in alpha-helix and beta-sheet conformation in the fully stretched structure ( $\approx 320$ ) is much smaller than its initial value ( $\approx 1130$ ). This is due to the fact that the beta-sheets observed in the stretched tetramer are composed of discrete short beta-structured segments with a characteristic length of  $\approx 28$  Å (corresponding to 7 amino acids in length, with roughly 7 H-bonds each), while for stretched dimers the average beta-sheet length is  $\approx 45$  Å. Interdimer sliding sets in at strains in excess of 180%, which is mediated by the initiation of rupture of the small beta-sheet protein crystals.

Figure 9 shows the structural spectrum of the vimentin dimer (Fig. 9A) and tetramer (Fig. 9B), as a function of strain, providing details into the structural transition of the protein as the strain is increased. The plot shows the structural character associated with each of the 466 amino acids, revealing the spatial distribution of the structural changes during deformation. The beta-sheets form a large number of relatively short segments in the overlapped region, leading to significant oscillations of the stretching force as shown in regime (IV) of Fig. 4B.

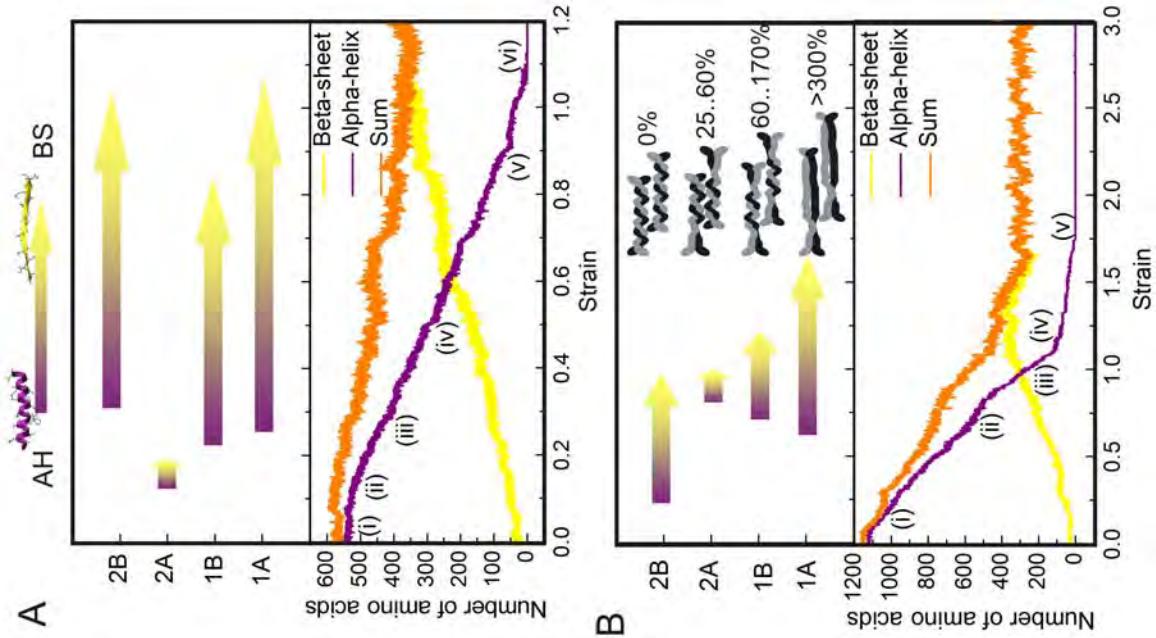
The deformation of IFs, involving the  $\alpha/\beta$  transition and subsequent sliding suggests that significant filament thinning should occur. Indeed, this behavior has been observed experimentally in atomic force microscopy and electron microscopy studies of desmin IFs [8,43] as shown in Fig. 10, in qualitative agreement with our predictions. Whereas the sliding of dimers or tetramers within a single filament has been proposed as a mechanism to explain the superelasticity of IFs, it was never directly observed at the single filament level. The only experimental evidence for sliding is based on a small angle x-ray diffraction study of stretched hard alpha-keratin fibers [44].

### Hierarchical structure is crucial to define mechanical properties

The hierarchical structure of IFs is crucial to achieve their characteristic mechanical properties, as it enables the cascaded activation of a different deformation mechanism, each at a specific level of tensile strain, to accommodate substantial mechanical deformation. Specifically, this cascaded activation of mechanisms at multiple levels in the structural makeup of IFs explains the remarkable behavior of IFs that renders them capable to withstanding extreme deformation and large loads. Thereby, the specific sequence of alpha-helix unfolding (providing great sensitivity to mechanical deformation, as shown in Figs. 8 and 9), subsequent  $\alpha/\beta$  transition and interdimer sliding via beta-sheet crystal rupture ensures the mechanical function of IFs under a wide range of strains. The key to enable these mechanisms is the existence of specific structures at each hierarchical level.

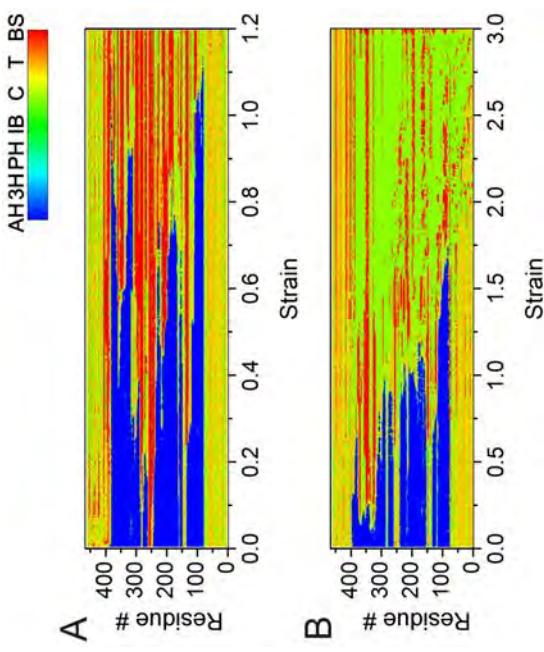
We discuss here the detailed role of each level of hierarchy in the filament. The unfolding of alpha-helical domains, facilitated by repeated breaking of serially arranged clusters of H-bonds (see Figs. 6 and 7) absorbs stretching energy while keeping the other structures intact. Thereby, the grouping of H-bonds into clusters of 3–4 in each alpha-helical turn has been shown to be crucial in order to yield maximum mechanical strength against rupture, and contributes to the enhanced stability of alpha-helical turns [35].

The coiled arrangement of the polypeptide into turns provides the structural basis for increased “hidden length” of the backbone chain, which is being released when rupture occurs. This, together with the serial arrangement of multiple turns into long alpha-helical domains is the structural basis to enable their ability to stretch more than 140% at relatively low force levels on the order of hundreds of pN. The arrangement of two alpha-helices into coiled-coil domains provides increased strength and additional



**Figure 8. Geometry-strain maps for the dimer (panel A) and tetramer (panel B) structures.** The lower part of the plots shows the number of amino acids associated with alpha-helical structure (purple) or beta-sheet structure (yellow), and the sum for these two known structures (orange). The arrows shown in the upper part of the plot denote processes of  $\alpha$ - $\beta$  transition for each segment. Each point marked corresponds to a snapshot with the same labels as shown in Figs. 6 and 7. The inlay in panel B shows schematics of the deformation mechanisms of the IF tetramer at various levels of strain.  
doi:10.1371/journal.pone.0007294.g008

deformability to increase the overall mechanical stretchability in excess of 150% [45]. During the uncoiling of alpha-helical and coiled-coil structures, an  $\alpha$ - $\beta$  transition occurs. This is facilitated by allowing the uncoiled chains to connect via H-bonds, and thereby to form betasheets (Fig. 6B), assembled into small clusters with several H-bonds each. These betasheet topologies are highly efficient in resisting tensile force and thus control the mechanical behavior in the regime of large deformation at strain levels in excess of 150% (see Figs. 8B and 9B) [46]. The betasheet shear topologies also facilitate a significant increase of the tangent modulus as they provide sufficient force resistance to induce stretching of the polypeptide backbone. Once the applied force



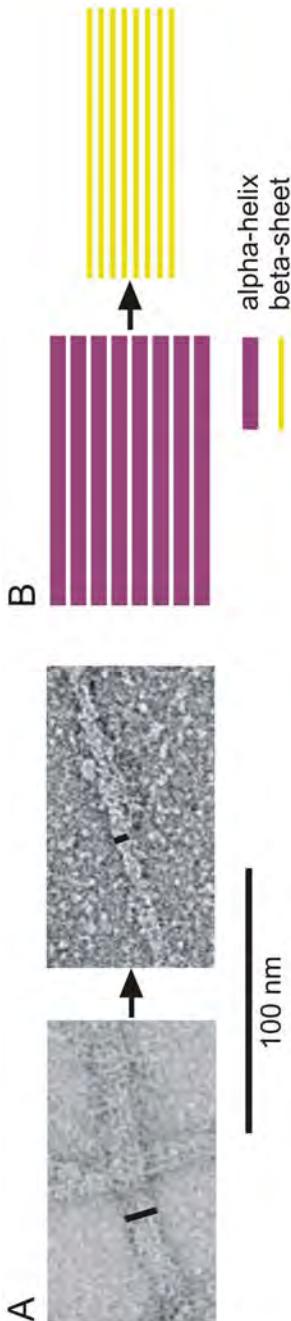
**Figure 9. Structural spectrum of the vimentin dimer and tetramer, as a function of strain.** The color bar indicates the structural character is separated into seven parts and corresponds to an  $\alpha$ -helix (AH), 3–10 helix (3H), phi-helix (PH), isolated-bridge (IB), coil (C), transition (T) and a  $\beta$ -sheet (BS). Panel A: The spectrum of the structural character of each amino acid of a monomer (with 466 amino acids in total) in the vimentin dimer during the entire stretching process (due to the symmetric arrangement of the other monomer they show a same behavior, and is thus not shown). Panel B: The spectrum of the structural character of each amino acid of a monomer (with 466 amino acids in total) in the vimentin tetramer during the entire stretching process (due to the symmetric and antisymmetric arrangement of the other three monomers, they show the same behavior, and are thus not shown). The analysis shows the gradual transition from an alpha-helix dominated structure (with unstructured and beta-sheet based linker/head/tail domains) towards a beta-sheet rich structure. In the case of the vimentin tetramer, the onset of sliding at  $\approx 170\%$  strain induces a reduction of the beta-sheet content as the strain is increased. The betasheets form many short segments in the overlapped region beyond this level of strain, leading to significant oscillations of the stretching force as shown in regime (IV) of Fig. 4B.  
doi:10.1371/journal.pone.0007294.g009

exceeds a critical value at 175% strain, interprotein sliding sets in as beta-sheets begin to rupture in a stick-slip like fashion, resulting in a force-plateau with relatively large forces that continues up to 300% strain when the proteins detach and catastrophic failure occurs. Altogether, these mechanisms explain how IFs can sustain 300% deformation without catastrophic failure, while providing a continuously increasing mechanical stiffening as the tensile deformation is increased.

Table 1 summarizes the function of each hierarchical level in vimentin IFs in defining their mechanical properties. Overall, this analysis suggests that the hierarchical makeup of IFs is crucial in defining their characteristic biomechanical properties.

## Discussion

Here we developed an atomistic model of human vimentin dimer and tetramer IFs, validated against key structural parameter identified in experiment (Fig. 3). We carried out molecular dynamics simulations to measure their response to mechanical stress. To the best of our knowledge, this is the first nanomechanical analysis of the vimentin dimer and tetramer at atomistic and molecular resolution. Our analysis provides detailed insight into



**Figure 10. Experimental evidence for interdimer sliding and subsequent filament thinning.** Panel A: Electron microscopy picture of a mutant IF desmin DsA360P filament before and after shearing (adapted from reference [43]). Significant thinning of the filament diameter of  $\approx 55\%$  seen in these experimental images suggests that  $\alpha\beta$  transitions and interfilament sliding occur, in agreement with simulation results (the diameter changes from  $\approx 12$  nm to  $\approx 5.5$  nm). Panel B: Schematic of the deformation process, illustrating the thinning of the filament due to formation of beta-sheets.

doi:10.1371/journal.pone.0007294.g010

the molecular-level deformation mechanisms, and enabled us to link the force-strain behavior (Fig. 4) to geometric changes in the structural makeup of the protein at different hierarchical levels (Figs. 6 and 7 and Table 1). Our results are validated against experimental studies of the mechanical properties of IFs (Fig. 4C and 5). The overall good agreement between experiment and simulation suggests that our model captures the key structural and mechanical features of the vimentin IF. Specifically, the tetramer molecular simulation study highlights the two molecular ingredients elementary for IF tensile properties, an alpha-helix to beta-sheet transition, and a sliding process that begins at large deformation.

The most salient feature of vimentin and other IFs is that the filament is rather compliant at low strain, whereas it becomes much stiffer at high strain, as confirmed in the analysis shown in Fig. 4. We find that there is no sliding at small strain, suggesting that IFs should be considered as a continuous body at small strains. This is confirmed by the fact that the predictions put forth based on the results shown in Fig. 4A agree closely with experimental modulus measurements in bending. We find that large deformation of both the dimer and the tetramer is accompanied by an  $\alpha\beta$  transition, a phenomenon that has been observed for other long

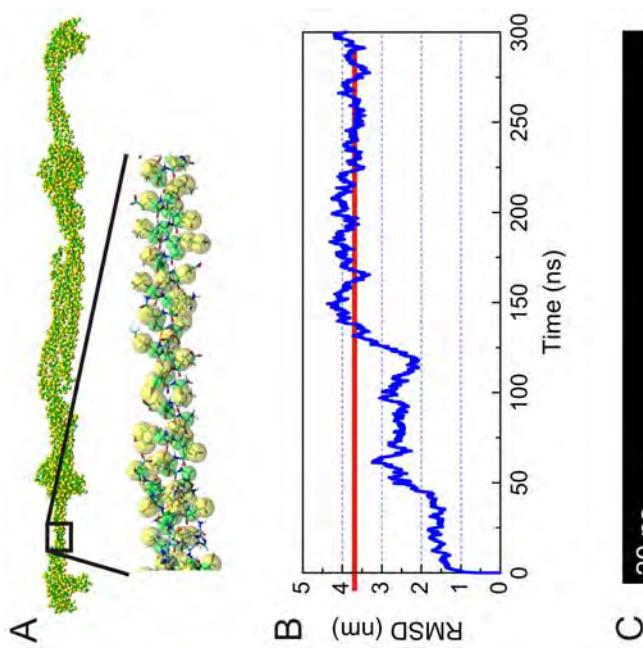
coiled-coil protein filaments such as  $\alpha$ -keratin IF fibers [39,41,42], myosin [37] and IF hagfish slime threads [6,34]. Our results further suggest that the hierarchical makeup of IFs is elementary in defining their characteristic mechanical properties through a cascaded activation of deformation mechanisms, each associated with a specific level of filament strain as shown in Table 1. This process, likely evolutionarily driven, combines disparate properties such as strength and superelasticity (the ability to tolerate large deformation without breaking). At larger levels, these properties may play a role in the protection of the structure and topology of the cell's cytoplasm under large deformation and at large forces. Furthermore, the characteristic sequence of segment unfolding as a function of different levels of strain (see e.g. Fig. 8) may be an important element in mechanotransduction, as each unfolded segment could perhaps facilitate the binding of specific signaling proteins, providing a means to enable biochemical signaling of different levels of stretch.

The assumptions made in the design of the computational approach used here are based on using methods that allow for the development of the most appropriate simulation model, given constraints such as simulation size, computational resources, and requirements based on the particular biophysics of the problem.

**Table 1. Role and mechanism of individual levels of structural hierarchies in IFs (see Fig. 1 for schematic of structure and labeling of hierarchical levels).**

Hierarchy level $H_n$	Description	Key mechanisms
H0	Level of chemistry: intrabackbone H-bond	H-bonds form at moderate temperatures; Drive formation of alpha-helices
H1	Alpha-helical turn defined by cluster of 3–4 H-bonds	Clusters of 3–4 H-bonds provide maximum mechanical strength at minimal material cost [35]; Coiled geometry of polypeptide provides significant “hidden length”
H2	Alpha-helix domain	Linear array of alpha-helical turns enables large extensibility of $>150\%$ strain via repeated rupture of turns [35,36]
H3	Coiled-coil protein domain	Increased stability and mechanical resistance [45]; Additional extensibility through coiled geometry [37]
H4	Dimer	Combination of several coiled-coil domains connected via linkers facilitates controlled unfolding segments (Fig. 6); Linkers form hinge-like structures (Fig. 2B)
H5	Tetramer	Staggered geometry provides basis for $\alpha\beta$ transition with increased resistance to interfilament shear through formation of beta-shear topologies; Sustain stretching up to $\approx 300\%$ at large force levels (Fig. 7)
H6	Unit length filament	Assembly intermediate (intermediate state between trimers and full length IF filaments)
H7	Full length IF filament	Axially assembled from UIFs. Filaments with great stretchability, stiffening and superelastic properties (Fig. 4C)
$>H7$	Cytoskeleton, cell	Not explored here

The cascaded activation of mechanisms at multiple hierarchical levels is the key to explain the unique properties of IFs, and illustrates that each level of hierarchy contributes an important element to the overall properties at the filament level. Thereby, the structure at each level controls the associated mechanism, and the aggregate effect of all mechanisms explains the characteristic mechanical properties of IFs.



**Figure 11. Coarse-grained representation of the vimentin tetramer (in explicit water) and structural stability of the vimentin tetramer structure at time-scales of hundreds of nanoseconds.** By replacing the full atomistic representation with a bead model (where several beads reflect one amino acid), much longer time-scales can be reached. Panel A shows the coarse-grained representation of the vimentin tetramer (water molecules not shown for clarity). Panel B depicts the root mean square displacement function as a function of time, showing structural stability of the predicted structure. The root mean square displacement function provides a measure for how much a structure changes relative to the initial geometry. Panel C shows snapshots during the coarse-grained equilibration, indicating the bead domains in a different color.  
doi:10.1371/journal.pone.0007294.g011

Future work could be focused on further refinement of the structural model reported here, for example by using higher fidelity atomistic force field models, or perhaps by using methods that facilitate the enhanced sampling of structural configurations. However, due to the large size of the protein filaments considered here, the computational expense of such simulations can be significant, in particular if long time-scales are considered.

The use of other methods, such as replica exchange simulations or coarse-grained approaches could be explored to provide an additional comparison with the results obtained here, for example to reach a better sampling of structural configurations at long time-scales. Here we report a first simulation of the vimentin tetramer using a coarse-grained model with explicit water, as shown in Fig. 11A, where the initial geometry is based on the result of the full atomistic equilibration (structure shown in Fig. 2C). This coarse-grained model enables us to reach time-scales of hundreds of nanoseconds to microseconds. An analysis of the structural evolution of the tetramer in this coarse-grained representation provides further evidence for the stability of the predicted structure at long time-scales of hundreds of nanoseconds, as confirmed in the analysis shown in Fig. 11B and Fig. 11C. Since the structure is stable with the atomistic (both effective Gaussian model and explicit solvent), as well as with the coarse-grained representation provides evidence that the results are reliable, despite variations in the model representation. The coarse-grained model may be applied to studying the assembly process of unit length filaments or the compacting process of full length filaments.

Our analysis illustrates the opportunities associated with a computational approach in describing complex biological materials. The study reported here provides a new way forward in carrying out molecular-level studies of IF structure-property relationships. This may be particularly important to probe the effects of the variability of the structure and associated changes in chemical and intermolecular interactions (e.g. amino acid mutations and other molecular defects). Defects may be introduced at different hierarchical levels of the structure of vimentin intermediate filament, and the effects on the biomechanical performance may provide important insight into physiological and diseased states of IFs. This specifically includes the possibility of probing the effect of mutations using *in silico* materials methods, enabling pathogenesis studies of severe medical conditions associated with IF. The model sets the stage for studies ranging from *in vivo* cross-linking events (e.g. how keratinocyte transglutaminase crosslinks keratins in the skin) to investigating the effects of protein binding on IF mechanics. Due to the emerging significance of intermediate filaments in a variety of cellular process, the results of our study also contribute to improved understanding of cell deformation, cell adhesion and mechanotransduction. IFs can also be considered a model system that may enable us to fabricate *de novo* engineered materials that display a high sensitivity to applied forces, show resilient mechanical properties, and provide biological compatibility. Another possible next challenge of using such models is to apply them to describe larger-scale cellular structures [47] and to develop bottom-up models of dynamic processes such as cell signaling and protein expression. For example, IF networks *in vivo* are in equilibrium with (and interact with) a pool of soluble IF associated proteins that are able to exchange anywhere along the filaments. With our approach, we may be able to address how moderate tension may change the exchange rate.

The concept of designing materials with hierarchical structures, by deliberately determining a cascade of multi-scale mechanisms as shown in Table 1 is a largely unexplored aspect in materials science that could lead to advances in *de novo* materials design. By

utilizing self-assembly processes from nano to macro, hierarchical structures may be the key that can enable us to take advantage of properties at all scales, and to exploit superior nanoscale properties. This could lead to a new class of biocompatible natural materials with applications as stimulus responsive gels, self-assembled nanomaterials, biomaterials, or as energy absorbing materials.

## Materials and Methods

### Atomistic level molecular models and force fields

Molecular simulations are carried out using the CHARMM19 all-atom energy function with an effective Gaussian model for the water solvent [22,23]. Additional simulations are carried out using explicit solvent models with the CHARMM force field and explicit TIP3 water [24] implemented in NAMD [25] for validation of the effective Gaussian simulations. In both models, we apply a 1 femtosecond time step, with constant temperature (300 K). A pressure of 1 atmosphere is used in the explicit water simulations. The implicit water molecular model is currently the only feasible approach to simulate sufficiently large protein structures at the atomistic level at long time-scales (reaching tens of nanoseconds), and is therefore used for structure prediction and mechanical characterization. The explicit solvent model and the coarse-grained approach is used to validate the results of implicit solvent simulations. The use of the effective solvent model is further motivated by our desire to overcome the limitations of relatively fast unfolding rates used in the simulations. The lack of slow water relaxation processes in the implicit model enables us to effectively model the physical conditions of pulling experiments at much longer time-scales. Therefore, effective solvent methods provide physically more meaningful results than the use of explicit solvent methods, as discussed in [48,49,50].

Several other studies have recently utilized the approach used here. For example, the effective Gaussian model for the water solvent was used in mechanical unfolding studies of proteins in [48] and [50], where mechanical properties of several types of protein segments were studied. Structure prediction approaches similar as used here have also been reported in [49], where the conformation of collagen structures has been studied. Due to the size of the protein, the simulations reported here take up to 50 days each on a parallel Linux cluster.

### Coarse-grained model and force field

In addition to atomistic level simulations we carry out an equilibration of the molecular structure by using the coarse-grained MARTINI force field [26] in explicit water, to explore the structural stability of the predicted structures at time-scales of several hundred nanoseconds. The initial coarse-grained geometry (shown in Fig. 11A) is generated from the full atomistic model of the tetramer (after the structure prediction process). We use a 40 femtosecond time step in this coarse-grained model, and the same constant temperature and pressure as in the full atomistic model. The coarse-grained system includes 43,801 particles (4,052 for the protein plus 39,749 for water molecules), while the corresponding full atomistic system would include 506,915 atoms (29,924 atoms for the protein plus 476,991 atoms for water molecules). The number of particles in the coarse-grained model is 1/11 of the full atomistic representation.

### Molecular model setup

The structure of the dimer is predicted by a series of computational steps as summarized in Fig. S1. First, the consideration of experimental sequence information and associated structural features enables us to develop an initial configuration that is close to the naturally favored state. Through this approach, we avoid solving a complete protein folding problem and rather focus on structural optimization close to the equilibrium geometry. Alpha-helical parts are bundled and rotated considering the orientation of their hydrophobic strikes. Based on the initial structure, equilibrated geometries are obtained through a series of repeated structural optimization steps, using both energy minimi-

zation and equilibration (following an approach suggested in [22,23]). Figure S2 shows the evolution of the total energy and the root mean square deviation (RMSD) for both structures, during the last 10 ns of the equilibration period. The convergence of the two curves at constant values suggests that both the dimer and the tetramer have reached a stable configuration.

### Validation of molecular model

To confirm the stability of the molecular structures, we carry out simulations in explicit solvent using the CHARMM force field with explicit solvent [24]. Figures S3 and S4 show the results of radial distribution function analyses, comparing the structures in implicit and explicit solvent for both the dimer and tetramer, confirming the structural stability of the molecular models. Furthermore, we compare the radial distribution function of our model with the existing segments obtained from x-ray diffraction (PDB entries 1gk7 for the 1A segment and 1gk4 for the 2B segment), showing good agreement as depicted in Fig. S5.

The atomistic geometry resulting from the structure prediction runs is also used to build a coarse-grained representation of the vimentin tetramer using the MARTINI force field [26], where an analysis of the stability of the molecule at hundreds of ns time-scales (Figure 11) and the analysis of the radial distribution function of the backbone particles (Fig. S6) suggests that the predicted geometry is stable.

Altogether, these comparisons and the structural data shown in Figure 3 and Table S1 suggests that the molecular models of the dimer and tetramer are reasonable representations of the vimentin IF dimer and tetramer structures.

### Nanomechanical characterization

The C<sup>α</sup> atoms at the end of two 2B segments are pulled on by using steered molecular dynamics, while the other end of the filament is fixed (with a force constant of 10 kcal/mol/Å<sup>2</sup>). The pulling force  $F$  is recorded versus the position. The simulations are carried out at pulling velocities ranging from 0.0001 Å/ps to 1 Å/ps. It is noted that 0.1 Å/ps is the lowest computationally accessible pulling speed for the larger tetramer system in which we can perform stretching until failure occurs. The analysis of the mechanical properties is facilitated by calculating engineering stress and strain. We record the force-displacement curve from steered molecular dynamics simulation and analyze the mechanical properties by computing the engineering stress and strain, which are defined as

$$\sigma = F/A_0, \quad (1)$$

$$\varepsilon = \Delta L/L_0, \quad (2)$$

where  $F$ ,  $A_0$ ,  $\Delta L$  and  $L_0$  are the pulling force, relevant cross-sectional area, displacement, and initial length, respectively. The (tangent) Young's modulus is determined by

$$E = \frac{\Delta\sigma}{\Delta\varepsilon}. \quad (3)$$

### Comparison between full-length filament experiment and simulations

The comparison between simulations of tetramers and experiment results of full length IFs as shown in Fig. 4C is done as

follows. In full length filaments there are no free ends. Therefore the structure is regarded as a rolled sheet, which is composed of eight overlapping parts. To reflect the lack of free ends in full length filaments we remove the part of the force-strain data that corresponds to the initial unfolding of the free ends. The modified strain corresponding solely to the overlapped part is obtained by converting the strain in the third and forth region in Fig. 4B following

$$z_{\text{overlap}} = (\varepsilon - \varepsilon_0) \frac{L_{\text{tetramer}}}{L_{\text{overlap}}} \quad (4)$$

(where  $\varepsilon_0$  is the beginning of the third region). We shift the simulation force to begin from zero at  $\varepsilon_0 = 60\%$ . To compute the engineering stress, we divide the force by the cross-sectional area of the tetramer,  $A_0 = 32.9 \text{ nm}^2$  (for details of geometric analysis, Please see Supporting Methods S1). Under changes in strain rate, the strength of protein filaments increase by a factor compared with the quasi-static loading observed in experiment [35,51]. To account for this effect, we calculate a force scaling factor to extrapolate the simulation results to the strain rate used in experiment, specifically considering the rate dependence of beta-sheet proteins (since the failure of beta-sheets accounts for the final peak in the stress-strain curve). Earlier studies have revealed that for beta-sheet proteins (under shear loading) the strength at  $0.1 \text{ \AA/pS}$  pulling speed is  $1,700 \text{ pN}$  [52]. For slow pulling on the order of  $10^{-11} \text{ \AA/pS}$  the strength is found to be  $240 \text{ pN}$  [53]. This leads to a scaling factor value of  $7.1$  to account for the different strain rates in experiment and simulation. This facilitates a direct comparison with experimental results using AFM [8].

### Molecular structure analysis

We use Visual Molecular Dynamics (VMD) [54] for visualization of protein structures as well as for the analysis of the  $\alpha$ - $\beta$  transition shown in Figures 8 and 9. The structure analysis is done using the approach suggested in [55]. The rupture length of H-bonds is  $3.7 \text{ \AA}$  for visualization in VMD. The criterion to identify H-bonding considers both hydrogen bond patterns and the backbone geometry, based on polypeptide chain dihedral angles.

### Supporting Information

**Methods S1** Additional information about computational methods, data analysis and research approach.  
Found at: doi:10.1371/journal.pone.0007294.s001 (0.10 MB PDF)

**Table S1** Structural parameters of the dimer/tetramer structure and a quantitative comparison between experiment and simulation. Part of the data shown here is visualized in Figure 3.  
Found at: doi:10.1371/journal.pone.0007294.s002 (0.01 MB PDF)

**Table S2** Mechanical properties of vimentin IF, quantitative comparison between experiment and simulation. Part of the data shown here is visualized in Figure 5.  
Found at: doi:10.1371/journal.pone.0007294.s003 (0.06 MB PDF)

**Figure S1** Overview over the structure prediction approach used here. The incorporation of structural features based on amino acid sequence is used to create an initial structural model. A sequence of energy minimization and equilibration, repeated until convergence is reached, results in the final structure that is validated against experimental results and then used for mechanical analysis.

Found at: doi:10.1371/journal.pone.0007294.s004 (0.03 MB PDF)

**Figure S2** Total energy and root mean square displacement (RMSD) analysis for the last 10 ns of the equilibration process, for the dimer (panel A) and the tetramer (panel B).  
Found at: doi:10.1371/journal.pone.0007294.s005 (0.03 MB PDF)

**Figure S3** Radial distribution function (RDF) for both models in implicit solvent and explicit solvent (panel A: dimer, panel B: tetramer). The peaks represent the distance from an alpha-carbon atom to the nearest neighbor alpha-carbon atoms, indicating the secondary and tertiary structure of coiled-coil proteins. The same location of the peaks means that structural characters are same for our protein model in both the implicit solvent and explicit solvent environment.  
Found at: doi:10.1371/journal.pone.0007294.s006 (0.03 MB PDF)

**Figure S4** Integrated of RDF function for both models in implicit solvent and explicit solvent (panel A: dimer, panel B: tetramer).  
Found at: doi:10.1371/journal.pone.0007294.s007 (0.03 MB PDF)

**Figure S5** Comparison of RDF analysis between our model and experimental results (based on the model obtained through x-ray diffraction analyses), for the 1A segment (panel A), and for the 2B segment (panel B). The peaks represent the distance from an alpha-carbon atom to the nearest neighbor alpha-carbon atoms, indicating the secondary and tertiary structure of coiled-coil proteins. The same location of the peaks means that structural characters are same for our protein model and experimental model.  
Found at: doi:10.1371/journal.pone.0007294.s008 (0.03 MB PDF)

**Figure S6** Comparison of the RDF between the full-atomistic model and the coarse-grained representation, after 300 ns equilibration. The peaks represent the distances from a backbone bead to the nearest neighbor backbone beads, indicating the secondary and tertiary structure of coiled-coil proteins. The same location of the peaks means that structural characters are same for our protein model in both the implicit solvent and explicit solvent environment.  
Found at: doi:10.1371/journal.pone.0007294.s009 (0.04 MB PDF)

**Structure S1** Atomistic structure of the intermediate filament dimer in the Protein Data Bank (PDB) format.  
Found at: doi:10.1371/journal.pone.0007294.s010 (0.74 MB TXT)

**Structure S2** Atomistic structure of the intermediate filament tetramer in the Protein Data Bank (PDB) format.  
Found at: doi:10.1371/journal.pone.0007294.s011 (1.47 MB TXT)

**Movie S1** Equilibrated structure of the vimentin IF dimer at 300 K. The movie shows a 5 ns interval of a constant temperature simulation of the dimer in water solvent.  
Found at: doi:10.1371/journal.pone.0007294.s012 (6.79 MB AVI)

**Movie S2** Equilibrated structure of the vimentin IF tetramer at 300 K. The movie shows a 5 ns interval of a constant temperature simulation of the tetramer in water solvent.  
Found at: doi:10.1371/journal.pone.0007294.s013 (8.36 MB AVI)

## Author Contributions

Conceived and designed the experiments: ZQ LK MJB. Performed the experiments: ZQ MJB. Analyzed the data: ZQ LK MJB. Contributed to the writing and editing of the manuscript: ZQ LK MJB. All authors read and approved the final manuscript.

## References

- Hermann H, Aebi U (2004) Intermediate filaments: Molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. *Annual Review of Biochemistry* 73: 749–789.
- Hutchison CJ (2002) Lamins: Building blocks or regulators of gene expression? *Nature Reviews Molecular Cell Biology* 3: 838–858.
- Hermann H, Bar H, Krepak L, Strelkov SV, Aebi U (2007) Intermediate filaments: from cell architecture to nanomechanics. *Nature Reviews Molecular Cell Biology* 8: 562–573.
- Wang N, Butler JP, Ingber DE (1993) Mechanotransduction across the cell surface and through the cytoskeleton. *Science* 260: 1124–1127.
- Wang N, Stamenovic D (2002) Mechanics of vimentin intermediate filaments. *Journal of Muscle Research and Cell Motility* 23: 535–540.
- Fudge D, Russell D, Beriault D, Moore W, Lane EB, et al. (2008) The intermediate filament network in cultured human keratinocytes is remarkably extensive and filamentous. *PLoS ONE* 3: e2327.
- Lewis MK, Nahinry PC, Chen V, Adhikari BB, Wright J, et al. (2003) Concentric intermediate filament lattice links to specialized Z-band junctional complexes in sonic muscle fibers of the type I male midshipman fish. *Journal Of Structural Biology* 143: 56–71.
- Krepak L, Hermann H, Aebi U (2008) Tensile properties of single desmin intermediate filaments. *Biophysical Journal* 94: 2790–2799.
- Dahl KN, Kahn SM, Wilson KL, Discher DE (2004) The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. *Journal of Cell Science* 117: 4779–4786.
- Wilson KL, Zastrow MS, Lee KK (2001) Lamins and disease: Insights into nuclear infrastructure. *Cell* 104: 647–650.
- Ackbarow T, Sen D, Thaulow C, Buehler MJ (2009) Alpha-Helical Protein Networks are Self Protective and Flaw Tolerant. *PLoS ONE* 4: e6015.
- Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, et al. (2001) Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. *Nature Genetics* 27: 117–120.
- Omari MB, Coulombe PA, McLean WH (2004) Intermediate filament proteins and their associated diseases. *N Engl J Med* 351: 2087–2100.
- Parry DAD, Strelkov SV, Burkhardt P, Aebi U, Hermann H (2007) Towards a molecular description of intermediate filament structure and assembly. *Experimental Cell Research* 313: 2204–2216.
- Sokolova AV, Krepak L, Wedig T, Mucke N, Svengut DI, et al. (2006) Monitoring intermediate filament assembly by small-angle x-ray scattering reveals the molecular architecture of assembly intermediates. *Proceedings of the National Academy of Sciences of the United States of America* 103: 16206–16211.
- Strelkov SV, Hermann H, Geisler N, Lusig A, Ivaninskii S, et al. (2001) Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. *Journal of Molecular Biology* 306: 773–781.
- Strelkov SV, Hermann H, Geisler N, Wedig T, Zimbelmann R, et al. (2002) Conserved segments IA and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. *Embo Journal* 21: 1255–1266.
- Rafik ME, Doucet J, Buehler MJ (2004) The intermediate filament architecture as determined by X-ray diffraction modeling of hard alpha-keratin. *Biophysical Journal* 86: 3893–3904.
- Luca S, Yau WM, Leipman R, Tycko R (2007) Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *Biochemistry* 46: 13505–13522.
- Goldie KN, Wedig T, Mira AK, Aebi U, Hermann H, et al. (2007) Dissecting the 3-D structure of vimentin intermediate filaments by cryo-electron tomography. *Journal Of Structural Biology* 158: 378–385.
- Qin Z, Granford S, Ackbarow T, Buehler MJ (2009) Robustness-strength performance of hierarchical alpha-helical protein filaments. *International Journal of Applied Mechanics* 1: 85–112.
- Lazaridis T, Karplus M (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations. *Science* 278: 1928–1931.
- Lazaridis T, Karplus M (1999) Effective energy function for proteins in solution. *Proteins-Structure Function and Genetics* 35: 133–152.
- MacKerell AD, Bashford D, Bellott M, Dunbrack RL, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal Of Physical Chemistry B* 102: 3586–3616.
- Nelson MT, Humphrey W, Gursay A, Dalko A, Kale LV, et al. (1996) NAMD: A parallel, object oriented molecular dynamics program. *International Journal Of Supercomputer Applications And High Performance Computing* 10: 251–268.
- Monticelli L, Kandasamy SK, Penole X, Larson RG, Tieleman DP, et al. (2008) The MARTINI coarse-grained force field: Extension to proteins. *Journal of Chemical Theory and Computation* 4: 819–834.
- Mucke N, Wedig T, Buer A, Marekov LN, Steinert PM, et al. (2004) Molecular and biophysical characterization of assembly-starter units of human vimentin. *Journal of Molecular Biology* 340: 97–114.
- Krepak L, Aebi U, Hermann H (2004) Molecular mechanisms underlying the assembly of intermediate filaments. *Experimental Cell Research* 301: 77–83.
- Smith TA, Strelkov SV, Burkhardt P, Aebi U, Parry DAD (2002) Sequence comparisons of intermediate filament chains: Evidence of a unique functional/structural role for coiled-coil segment Ia and linker Ia. *Journal of Structural Biology* 137: 128–145.
- Brown JH, Cohen C, Parry DAD (1996) Hepatid breaks in alpha-helical coiled coils: Stutters and stammers. *Proteins-Structure Function and Genetics* 26: 134–145.
- Hermann H, Haner M, Brettel M, Muller SA, Goldie KN, et al. (1996) Structure and assembly properties of the intermediate filament protein vimentin: The role of its head, rod and tail domains. *Journal of Molecular Biology* 264: 933–953.
- Parry DAD, Steinert PM (1999) Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism. *Quarterly Reviews of Biophysics* 32: 99–187.
- Hermann H, Haner M, Brettel M, Ku NO, Aebi U (1999) Characterization of distinct early assembly units of different intermediate filament proteins. *Journal of Molecular Biology* 286: 1403–1420.
- Fudge DS, Gardner KH, Forsyth VT, Riekel C, Gosline JM (2003) The mechanical properties of hydrated intermediate filaments: Insights from hagfish slime threads. *Biophysical Journal* 85: 2015–2027.
- Ackbarow T, Chen X, Ketten S, Buehler MJ (2007) Hierarchies, multiple energy barriers, and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains. *Proceedings of the National Academy of Sciences of the United States of America* 104: 16410–16415.
- Ackbarow T, Buehler MJ (2007) Super-tight, energy dissipation and strain hardening of vimentin coiled-coil intermediate filaments: Atomistic and continuum studies. *Journal of Materials Science* 42: 8771–8787.
- Schwaiger I, Sattler C, Hostetter DR, Rieff M (2002) The myosin coiled-coil is a truly elastic protein structure. *Nature Materials* 1: 232–235.
- Guzman C, Jeney S, Krepak S, Kulik AJ, et al. (2006) Exploring the mechanical properties of single vimentin intermediate filaments by atomic force microscopy. *Journal of Molecular Biology* 360: 623–630.
- Hearle JWS (2000) A critical review of the structural mechanics of wool and hair fibres. *International Journal of Biological Macromolecules* 27: 129–138.
- Microtubules and Actin-Filaments Measured From Thermal Fluctuations in Shape. *Journal of Cell Biology* 120: 923–934.
- Krepak L, Doucet J, Buehler MJ (2001) Unraveling double stranded alpha-helical coiled coils: An x-ray diffraction study on hard alpha-keratin fibers. *Biopolymers* 58: 526–533.
- Krepak L, Doucet J, Dumas P, Biki F (2004) New aspects of the alpha-helix to beta-sheet transition in stretched hard alpha-keratin fibers. *Biophysical Journal* 87: 640–647.
- Krepak L, Bar H (2009) Severe Myopathy Mutations Modify the Nanomechanics of Desmin Intermediate Filaments. *Journal Of Molecular Biology* 385: 1043–1051.
- Krepak L, Franbourg A, Biki F, Leroy F, Dalle D, et al. (2002) A new deformation model of hard alpha-keratin fibers at the nanometer scale: implications for hard alpha-keratin intermediate filament mechanical properties. *Biophys J* 82: 2265–2274.
- Krepak L, Schwartz R (2007) Computational models of stutter defects in vimentin intermediate filaments. *Experimental Mechanics* 49: 79–89.
- Keten S, Buehler MJ (2008) Geometric Confinement Governs the Rupture Strength of H-Bond Assemblies at a Critical Length Scale. *Nano Letters* 8: 743–748.
- Ackbarow T, Buehler MJ (2009) Molecular mechanics of stutter defects in organization in cellular environments. *Cell Biochem Biophys* 48: 16–31.
- Paci E, Karplus M (2000) Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc Natl Acad Sci U S A* 97: 6521–6526.
- Stultz CM (2006) The folding mechanism of collagen-like model peptides explored through detailed molecular simulations. *Protein Science* 15: 2166–2177.
- West DK, Brockwell DJ, Ohmsted PD, Radford SE, Paci E (2006) Mechanical resistance of proteins explained using simple molecular models. *Biophysical Journal* 90: 287–297.
- Sotomayor M, Schulten K (2007) Single-molecule experiments *in vitro* and *in silico*. *Science* 316: 1144–1148.
- Keten S, Buehler MJ (2008) Asymptotic strength limit of hydrogen-bond assemblies in proteins at vanishing pulling rates. *Physical Review Letters* 100: 198301.
- Sulkowska JI, Cieplak M (2007) Mechanical stretching of proteins - a theoretical survey of the Protein Data Bank. *Journal of Physics-Condensed Matter* 19: 283201.

54. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics.  
Journal Of Molecular Graphics 14: 33.

55. Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. *Proteins-Structure Function and Genetics* 23: 566-579.

# Molecular Dynamics Simulation of the $\alpha$ -Helix to $\beta$ -Sheet Transition in Coiled Protein Filaments: Evidence for a Critical Filament Length Scale

Zhao Qin and Markus J. Buehler\*

*Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue Room 1-235A&B, Cambridge, Massachusetts, 02139, USA*

(Received 27 September 2009; published 12 May 2010)

The alpha-helix to beta-sheet transition ( $\alpha$ - $\beta$  transition) is a universal deformation mechanism in alpha-helix rich protein materials such as wool, hair, hoof, and cellular proteins. Through a combination of molecular and theoretical modeling, we examine the behavior of alpha-helical coiled-coil proteins with varying lengths under stretch. We find that the occurrence of the  $\alpha$ - $\beta$  transition is controlled by the length of constituting alpha-helical proteins. In the asymptotic limit, short proteins with less than 26 amino acids or 3.8 nm length reveal interprotein sliding, whereas proteins with greater lengths feature an  $\alpha$ - $\beta$  transition, leading to a significant increase in the protein's stiffness, strength, and energy dissipation capacity at large deformation. Our study elucidates the fundamental physics of this mechanism and explains why the  $\alpha$ - $\beta$  transition typically occurs in protein filaments with long alpha-helical domains.

DOI: 10.1103/PhysRevLett.104.198304

PACS numbers: 82.37.Rs, 62.20.M-, 62.25.-g, 87.15.bd

Alpha helices and beta sheets are major secondary structural motifs that organize the three-dimensional structure of proteins [1]. Whereas the sequence of amino acids defines a protein's specific secondary and higher level structures [2], the transition between alpha helices and beta sheets ( $\alpha$ - $\beta$  transition) is observed universally for a broad range of alpha-helix based protein filaments. The  $\alpha$ - $\beta$  transition was observed under variations of pH [3], temperature changes [4], and solvent composition alterations [5], and under mechanical deformation as shown by experiment and simulation [6–10]. Many alpha-helix rich proteins are crucial for structural support, in mechano-transduction and cell motility, and in defining the cell's stretchiness. Experimental studies revealed that  $\alpha$  keratin [6,7] (found in wool, hair, and hooves), hagfish slime threads [8], desmin [9], and vimentin [10] all display a  $\alpha$ - $\beta$  transition under deformation. The transition significantly thins filaments, but toughens and strengthens it at large deformation [10]. Furthermore, the  $\alpha$ - $\beta$  transition may also have important implications for diseases related to amyloid formation (s.a. Alzheimer's, Type II Diabetes, Parkinson's), where native proteins convert to beta-sheet rich proteins [11,12].

The experimental observation of this phenomenon for a variety of proteins suggests that the  $\alpha$ - $\beta$  transition under stretching is a universal mechanism of deformation, and as such is crucial for providing a link between a protein material's hierarchical structure and its mechanical response [13]. The understanding of the  $\alpha$ - $\beta$  transition in protein materials may be similarly important as studies of dislocation nucleation and propagation in metals or ceramics, which present the fundamental unit deformation mechanism in crystalline materials. However, compared to dislocations, the understanding of deformation mechanisms in protein materials remain poorly understood.

What is the molecular mechanism of the  $\alpha$ - $\beta$  transition, and is there a key structural parameter that controls it? Here, we report a systematic study to address this issue.

We consider 13 randomly picked alpha-helical coiled-coil proteins selected from the protein data bank (PDB) [14] with varying lengths, all subjected to mechanical stretching (for information on coiled-coiled proteins, see [15]). Each protein is first energy minimized [16] and then equilibrated at 300 K. All simulations are performed using an *NVT* ensemble (Nose-Hoover thermostat 1 fs time step, CHARMM19 all-atom energy function and an effective Gaussian model for the water solvent [17,18] to facilitate rapid sampling of configurations under loading [19]). After equilibration, each structure is stretched by steered molecular dynamics (SMD), following the shear loading [Fig. 1(a)]. Thereby, the  $C^\alpha$  atom at the left end of one helical strand is fixed to resemble the attachment at a substrate, while the  $C^\alpha$  atom at the right end of the other strand is linked to a harmonic spring, spring constant 10 kcal/mol/ $\text{Å}^2$ . We pull at a constant velocity of 0.01  $\text{Å}/\text{ps}$  along the axial direction [Fig. 1(a)], and the displacement  $\Delta L$  and pulling force are recorded. The tensile strain is expressed as  $\varepsilon = \Delta L/L_0$  where  $L_0$  is the initial end-to-end length of the protein [20].

By analyzing both the H-bond patterns and the backbone geometry with the STRIDE algorithm [21], we assess the secondary structure and associated changes during deformation. For each protein, the stretching simulations are repeated 10 times with different random seeds for the initial velocity distribution for good statistics. Based on the results of an ensemble of 10 simulations for each protein, we then calculate the percentage of the occurrence of the  $\alpha$ - $\beta$  transition as a function of the coiled-coil length.

We begin by considering a very short coiled coil and

a very long one (short coiled coil PDB ID 1ajy,

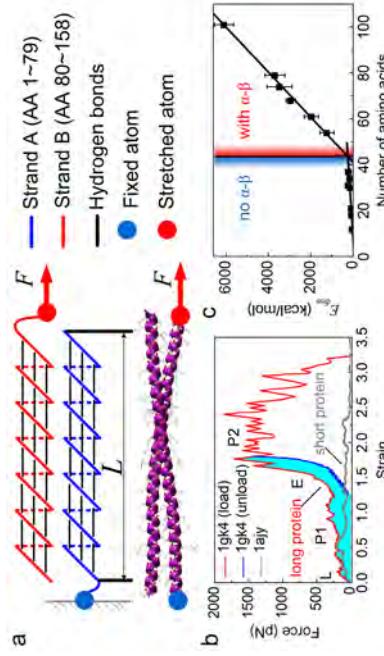


FIG. 1 (color online). Model for studying the strength of coiled-coil proteins under shear loading. (a) Upper: Schematic of the structure and loading condition of the coiled-coil ( $AA =$  amino acid residue). Lower: Example molecular structure of the 2B segment in the vimentin protein (PDB ID 1gk4). (b) Force-strain behavior of two coiled-coil proteins with different lengths (PDB ID 1gk4 (load), 1gk4 (unload), 1aj). In the long protein (1gk4), the modulus in the  $L$  region is 0.3 GPa average diameter 2.4 nm, and in the  $E$  region reaches 8.9 GPa average diameter 1.1 nm; PDB ID 1aj, short). The shaded area reflects the energy dissipated during unfolding. (c) Dissipated energy during the entire stretching process, for coiled-coil proteins with different lengths.

$L_0 \approx 2.6$  nm, long coiled coil 1gk4,  $L_0 \approx 11.5$  nm). In the short coiled coil, we see that the stretching force remains at a low value [Fig. 1(b)], where the maximum force approaches 200 pN at 55% strain as the two helical strands slide against each other [Fig. 2(a)]. We observe that only a few amino acids close to the ends of the protein domain are unfolded during the simulation [Fig. 2(c)]. In the long coiled coil, even though all simulation parameters are identical, the force-strain relation shows a dramatically different behavior compared with the short one, and now displays four phases [Fig. 1(b)]. First, we find a linear rise [phase ( $L$ )] towards a plateau at 200 pN [phase ( $P1$ )]. At the end of phase ( $P1$ ), we observe a regime with significant stiffening [phase ( $E$ )], followed by another, bumpier plateau phase ( $P2$ ) at relatively large force levels close to 1500 pN. A structural analysis in Fig. 2(d) suggests that the secondary structure remains intact within phase ( $L$ ), indicating that the protein undergoes solely elastic deformation below 15% strain (no H-bond rupture). The snapshots shown in Fig. 2(b) reveal that within phase ( $P1$ ), each strand twists as it unravels, and that groups of H bonds along the alpha-helix rupture, causing a sudden unfolding of helical turns. (Up until this point, the coiled coil behaves similarly as myosin [22,23].) Phase ( $E$ ) begins at force levels of 360 pN, but the force increases to more than 1000 pN at strains beyond 150%. In phase ( $P2$ ), a large number of beta sheets have formed, and an increase in the strain leads to the sliding of the two strands against one another, at a force level  $\approx 1500$  pN (1200–1800 pN; fluctuations due to the stretching, rupture, and rebinding

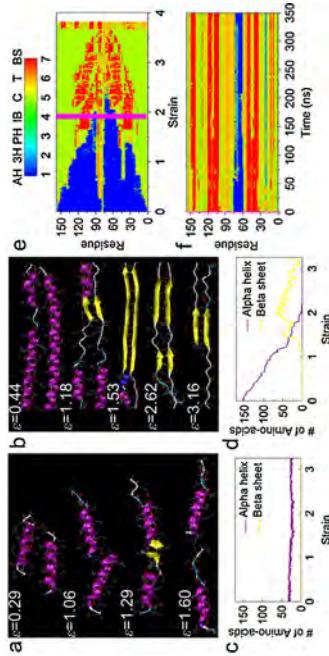


FIG. 2 (color online). Simulation snapshots of two coiled-coil proteins under shear (proteins with PDB ID 1ajy [(a), (c)] and 1gk4 [(b), (d)]). (a) and (b) Structure under applied strain during stretching of the proteins (H bonds are shown as dashed lines). (c) and (d) Number of amino acids associated with alpha-helix and beta-sheet secondary structures as a function of strain. (e) Structure spectrum for structure of the protein 1gk4 under stretching. (f) Structure spectrum of the protein 1gk4 during equilibration. The color bar distinguishes 7 structures: Alpha helix (AH), 3-10 helix (3H), Phi helix (PH), Isolated bridge (IB), Coil (C), Turn (T), beta sheet (BS).

of the H bonds in the sheared beta-sheet segments resembling a stick-slip motion). We note that this force level agrees with the strength of natural beta sheets reported earlier [24]. The much greater shear resistance is due to the fact that beta sheets in shear display a significantly larger strength [24–26] than alpha-helical proteins [22,27]. Phase ( $P2$ ) lasts until 325%, where the structure is completely detached. Figure 2(e) shows an analysis of the secondary structure as a function of strain, providing additional details into structural changes. This transformation of the secondary structure from an alpha-helix to a beta-sheet rich structure resembles the  $\alpha$ - $\beta$  transition observed in experiments.

We record the unloading  $F$ - $\varepsilon$  curve after the  $\alpha$ - $\beta$  transition has occurred. To achieve this, we begin from the configuration obtained at a strain of 180%, and subsequently relax the applied force to zero, then set the two ends free and equilibrate the structure. During 350 ns of equilibration, no structural changes are observed as confirmed by the spectrum analysis [Fig. 1(b)]. This irreversible character of the transition is also found in experimentally observed  $\alpha$ - $\beta$  transitions [7,10]. The unloading curve is depicted in Fig. 1(b), showing that the protein filament retracts during relaxation. However, the protein does not approach the length at the beginning of the simulation, suggesting permanent deformation. The shaded area in Fig. 1(b) represents the dissipated energy during the unfolding-refolding process [dissipated energy  $E_{\text{diss}} = \int_0^{180\%} (F_{\text{load}} - F_{\text{unload}}) L_0 d\varepsilon$ ]. Since the number of broken H bonds is 846 kcal/mol, calculated as  $E_{\text{diss}} = 846 (F_{\text{load}} - F_{\text{unload}}) L_0 d\varepsilon$ . Since the number of broken H bonds is  $N_H = 119$  at 180% strain [compared with the initial structure, Fig. 1(a)], we estimate the energy associated with each H bond as  $E_{\text{HB}} = E_{\text{tot}}/N_H = 7.11$  kcal/mol.

The initiation of the  $\alpha$ - $\beta$  transition process features three major phases. From phase (L) onwards, a single turn in the middle of the intact strand unravels locally inside one of the alpha helices (this explains why the ultimate force in this phase approaches the unfolding force of a single alpha helix). This structural defect, locally highly flexible, forms a nucleation point from which on nearby amino acids begin to unfold, turning them into random coils, as shown in Fig. 2(b) at  $\varepsilon = 0.44$  (the increase in flexibility can be explained by the significant decrease in persistence length, from several tens of nanometers for an alpha helix to approximately 0.4 nm for a random coil). The increasing number of unfolded turns in both helices increases the probability for hydrophobic interactions to take effect, which squeezes two unfolded turns together to form a first parallel beta-sheet cluster [Fig. 2(a),  $\varepsilon = 1.29$ ; Fig. 2(b),  $\varepsilon = 1.18$ ], here referred to as a *beta-sheet seed*. In the shorter coiled coil, there is either no beta-sheet seed or the beta-sheet seed is not strong enough to induce further unfolding of alpha-helical turns, and rather ruptures, leading to interprotein sliding [Figs. 2(a) and 2(c)]. In the longer coiled coil, however, the beta-sheet seed acts as an effective clamp that prevents sliding and thereby enforces unfolding of all other alpha-helical turns [Fig. 2(b),  $\varepsilon = 1.53$ ], resulting in the occurrence of the transition process throughout the rest of the chain [Fig. 2(d)]. At the end of the transition process, the structure is composed of an array of many small beta-sheet clusters [Fig. 2(b),  $\varepsilon = 2.62$ ]. The key to understand this process is to consider the ratio of the strength of the beta-sheet seed ( $f_B$ ) versus the strength of the alpha helix ( $f_A$ ), where the ratio of  $f_B/f_A$  controls whether ( $f_B/f_A \geq 1$ ) or not ( $f_B/f_A < 1$ ) the  $\alpha$ - $\beta$  transition occurs.

The off rate of a reference structure (alpha helix or beta sheet) with a reference length  $N_0$  is [27–29]:  $\chi_0 = \omega_0 \exp[-(E_b - F_0 x_b)/(k_B T)]$ , where  $\omega_0 = 10^{13} \text{ s}^{-1}$  is the natural frequency of bond vibration [30],  $x_b$  the unfolding transition point,  $k_B$  the Boltzmann constant,  $E_b$  the energy barrier,  $F_0$  the averaged unfolding force, and  $T$  the temperature. The off rate of alpha helices of general length  $N$  is  $\chi = \omega_0 (N/N_0) \exp[-(E_b - f_A x_b)/(k_B T)]$ , which considers the fact that alpha-helical turns are arranged in series (so that increasing the length increases the probability of unfolding). Noting the fact that pulling speed  $v = \chi x_b$  is constant, we obtain

$$f_A(N) = F_0 - \frac{k_B T}{x_b} \ln\left(\frac{N}{N_0}\right) \sim -\ln(N). \quad (1)$$

We note that longer alpha helices result in a reduced mechanical strength with a negative logarithmic dependence on  $N$ , a finding that has been validated in experiments and simulations [31]. We now calculate the strength of the beta-sheet seed,  $f_B$ . The energy barrier in the beta-sheet shearing is  $E_b = E_{\text{HB}} n_{\text{cr}}$ , where  $n_{\text{cr}} = 3–4$  is the number of H bonds that are involved in a unit rupture event

[32]. Combining with the expression of  $\chi_0$ , we arrive at

$$f_B(n_{\text{cr}}) = \frac{1}{x_B} \left[ k_B T \ln\left(\frac{v}{\omega_0 x_B}\right) + E_{\text{HB}} n_{\text{cr}} \right] \quad (2)$$

where  $v$  is the pulling speed and  $x_B$  the rupture distance along the shearing direction. Since the size of the beta-sheet seed is constant at  $n_{\text{cr}}$  regardless of the length of the alpha-helical proteins in the coiled coil,  $f_B(n_{\text{cr}}) = f_B = \text{const.}$

We now identify numerical values of the parameters in Eqs. (1) and (2) directly from the atomistic model. For Eq. (1), the strength of an alpha-helical protein with  $N_0 = 79$  at 0.01 Å/ps pulling speed is  $F_0 = 200 \text{ pN}$ , and  $x_b = 0.2 \text{ \AA}$  (parameters extracted from a pulling experiment of a single alpha-helical protein [27]). For Eq. (2),  $E_{\text{HB}} = 7.11 \text{ kcal/mol}$  (close to the value predicted from density functional theory [33]), and  $x_B = 4 \text{ \AA}$  [25]. As shown in Fig. 3(a),  $f_B$  of an initial beta-sheet seed falls within the range of  $346.5 \pm 61.5 \text{ pN}$ , whereas  $f_A$  varies greatly as a function of the length. The condition  $f_B = f_A(N_{\text{cr}})$  yields the critical amino acid number  $N_{\text{cr}} = 41 \pm 12$  from which on the  $\alpha$ - $\beta$  transition occurs.

We carry out simulations for 13 different coiled-coil proteins (corresponding PDB IDs shown in Fig. 3(b)), and measure the probability for the occurrence of the  $\alpha$ - $\beta$  transition. The results reveal that the points fall in two distinct regions and that the transition occurs at a critical number of  $N_{\text{cr}} = 38.7 \text{ amino acids (5.7 nm)}$ , which agrees with the theoretical analysis, thus corroborating the concept of a critical length for the  $\alpha$ - $\beta$  transition. It is noted that the outcome of the test for each protein is very robust. However, there are greater fluctuations for those proteins near the critical length (for these, the repeated simulations give appropriate statistics and convergence).

We find that the critical length decreases slightly under extreme variations of the pulling speed. For example, at  $v = 0.005 \text{ \AA/ps}$ , we obtain  $F_0 = 167 \text{ pN}$ ,  $f_B = 340 \text{ pN}$ , and  $N_{\text{cr}} = 34$ ; at  $v = 0.002 \text{ \AA/ps}$ ,  $F_0 = 146 \text{ pN}$ ,  $f_B = 330 \text{ pN}$ , and  $N_{\text{cr}} = 32$ ; at  $v = 0.001 \text{ \AA/ps}$ ,  $F_0 = 116 \text{ pN}$ ,  $f_B = 323 \text{ pN}$ , and  $N_{\text{cr}} = 29$ . At vanishing pulling velocities corresponding to experimental and physiological conditions (at 1  $\mu\text{m/s}$  and less), we estimate a critical number of 26 amino acids, or 3.8 nm ( $F_0 = 14 \text{ pN}$ ,  $N_0 = 288$ ,  $x_b = 1.2$ ,  $f_B = 96 \text{ pN}$ , and  $N_{\text{cr}} = 26$  based on work [25–27,34] [Fig. 3(c)].

Because of the distinct force-strain behavior, the critical length has major implications for the energy dissipation capacity of a protein filament [Fig. 1(c)]. Notably, energy dissipation increases manifold with the protein's length in the second regime where the  $\alpha$ - $\beta$  transition occurs. This finding provides a potential explanation for the universality of *long* alpha-helical protein filaments in mechanically relevant proteins in biology, as they provide an innate

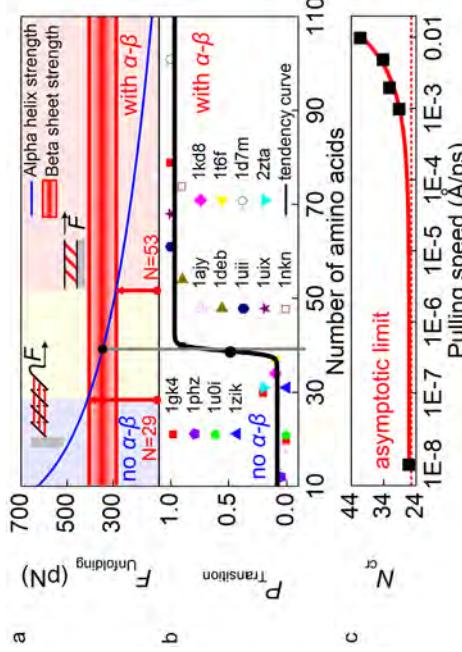


FIG. 3 (color online). (a) Theoretical analysis of the critical condition for the  $\alpha$ - $\beta$  transition, where the shaded area in the middle represents the critical condition. The left side of the critical point corresponds to sliding and unfolding of the alpha-helices without occurrence of the  $\alpha$ - $\beta$  transition, while the right side corresponds to the regime in which the  $\alpha$ - $\beta$  transition occurs. (b) Percentage of  $\alpha$ - $\beta$  transition occurrence for various coiled-coil proteins (simulation). Each point reflects the percentage of the transition within the simulation ensemble. The fitted curve is a sigmoid function  $p = A_2 + (A_1 - A_2)\{1 + \exp[-(n_0/dx)]\}^{-1}$ , with  $A_1 = 0.083$ ,  $A_2 = 0.967$ ,  $n_0 = 38.835$ , and  $dx = 0.465$ . The fit enables us to identify the transition point from the simulation data, leading to  $N_{cr} = 38.7$  (probability 50%). (c) Critical number of amino acids  $N_{cr}$  as a function of pulling speed (continuous curve) is power law fit to the simulation data: asymptotic limit  $N_{cr} = 26$ .

- [3] T. Koga *et al.*, *Chem. Eur. J.* **9**, 1146 (2003).
- [4] F. Ding *et al.*, *Proteins: Struct. Funct. Genet.* **53**, 220 (2003).
- [5] M. Meier and J. Seelig, *J. Am. Chem. Soc.* **130**, 1017 (2008).
- [6] J. S. Church, G. L. Corino, and A. I. Woodhead, *J. Mol. Struct.* **440**, 15 (1998).
- [7] L. Kreplak *et al.*, *Biophys. J.* **87**, 640 (2004).
- [8] D. S. Fudge *et al.*, *Biophys. J.* **85**, 2015 (2003).
- [9] L. Kreplak *et al.*, *Biophys. J.* **94**, 2790 (2008).
- [10] Z. Qin *et al.*, *PLoS ONE* **4**, e7294 (2009).
- [11] M. Gross, *Current Protein & Peptide Science* **1**, 339 (2000).
- [12] I. Daidone *et al.*, *Proteins: Struct. Funct. Bioinf.* **57**, 198 (2004).
- [13] M. J. Buehler and Y. C. Yung, *Nature Mater.* **8**, 175 (2009).
- [14] H. M. Berman *et al.*, *Nucleic Acids Res.* **28**, 235 (2000).
- [15] Coiled coils are double-stranded protein motifs (e.g., found in myosin and kinesin), where each strand is an alpha-helix with heptad repeated substrings. This seven residue repeat generally has apolar residues at the first and fourth position, forming a left-handed hydrophobic stripe. Following the orientation of this stripe, the two alpha-helical strands wrap around each other.
- [16] We minimize the potential energy via steepest descent and adopted basis Newton-Raphson (ABNR) methods.
- [17] T. Lazaridis and M. Karplus, *Science* **278**, 1928 (1997).
- [18] T. Lazaridis and M. Karplus, *Proteins: Struct. Funct. Genet.* **35**, 133 (1999).
- [19] E. Paci and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6521 (2000).
- [20] We use visual molecular dynamics (VMD), W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996), for visualization. The rupture length of H bonds is defined to be 5 Å for visualization purposes.
- [21] D. Frishman and P. Argos, *Proteins: Struct. Funct. Genet.* **23**, 566 (1995).
- [22] I. Schwaiger *et al.*, *Nature Mater.* **1**, 232 (2002).
- [23] D. D. Root *et al.*, *Biophys. J.* **90**, 2852 (2006).
- [24] M. Sotomayor and K. Schulten, *Science* **316**, 1144 (2007).
- [25] S. Keten and M. J. Buehler, *Nano Lett.* **8**, 743 (2008).
- [26] S. Keten and M. J. Buehler, *Phys. Rev. Lett.* **100**, 198301 (2008).
- [27] T. Ackbarow *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16410 (2007).
- [28] M. Rief *et al.*, *Phys. Rev. Lett.* **81**, 4764 (1998).
- [29] E. Evans, *Annu. Rev. Biophys. Biomol. Struct.* **30**, 105 (2001).
- [30] G. I. Bell, *Science* **200**, 618 (1978).
- [31] H. Dietz and M. Rief, *Phys. Rev. Lett.* **100**, 098101 (2008).
- [32] We find that the initial beta-sheet seed formed under shearing includes 3–4 H bonds, corresponding to the number of H bonds in an alpha-helical turn [27].
- [33] K. Tsemekhman *et al.*, *Protein Sci.* **16**, 761 (2007).
- [34] Z. Qin *et al.*, *Nanotechnology* **20**, 425101 (2009).
- [35] M. de Leeuw *et al.*, *PLoS ONE* **4**, e7296 (2009).
- [36] R. Burioni *et al.*, *Proteins: Struct. Funct. Genet.* **55**, 529 (2004).

capacity to heightened energy dissipation at large deformations

Our results provide a critical geometrical condition of the  $\alpha$ - $\beta$  transition and associated molecular-level effects. The concept put forth in our simple model may be used to explain other structural transitions, stability, and flexibility in proteins or polymers, and as such, may find future applications in bottom-up materials design [35,36]. The insight into the critical length scales may have applications in the design of novel peptide-based materials with high stiffness, novel hierarchical fibers, and materials with high extreme absorption capacity.

Support from AFOSR (Grant No. FA9550081-0321) is acknowledged.

1

Author: mhuehler@N

H. Toozé *Introduction*

New York 1999)

Kw 10K, 1999).

U. S. R. III, Natural

\*Corresponding author: mbuehler@MIT.EDU

[1] C. Branden and J. Tooze, *Introduction to Protein Structure*

(Garland Pub., New York, 1999).

[2] D.L. Minor and P.S. Kim, *Nature* (London) **367**, 660

(1994).

# Nanomechanics of functional and pathological amyloid materials

Tuomas P. J. Knowles<sup>1\*</sup> and Markus J. Buehler<sup>2,3,4\*</sup>

**Amyloid or amyloid-like fibrils represent a general class of nanomaterials that can be formed from many different peptides and proteins. Although these structures have an important role in neurodegenerative disorders, amyloid materials have also been exploited for functional purposes by organisms ranging from bacteria to mammals. Here we review the functional and pathological roles of amyloid materials and discuss how they can be linked back to their nanoscale origins in the structure and nanomechanics of these materials. We focus on insights both from experiments and simulations, and discuss how comparisons between functional protein filaments and structures that are assembled abnormally can shed light on the fundamental material selection criteria that lead to evolutionary bias in multiscale material design in nature.**

**A** wide range of natural and artificial peptides and proteins possess an intrinsic propensity to self-assemble into fibrillar nanostructures that are rich in  $\beta$ -sheet secondary structure. These fibrils are generally organized in a similar manner at the molecular level; they are characterized by  $\beta$ -strands that are oriented perpendicularly to the fibril axis, and connected through a dense hydrogen-bonding network, which results in supramolecular  $\beta$ -sheets that often extend continuously over thousands of molecular units<sup>1–5</sup> (Fig. 1). Such fibrillar structures first received attention through their association with diseases related to protein misfolding<sup>6</sup>, including neurodegenerative disorders such as Alzheimer's and Parkinson's diseases<sup>7,8</sup>, where normally soluble proteins are deposited pathologically into obdurate aggregates known as amyloid fibrils<sup>9–14</sup> (see Table 1 for an overview of key terminology and concepts). Subsequently, however, functional amyloid-like materials were discovered in varying roles throughout nature<sup>15–20</sup>. Moreover, under conditions where the natively folded states of proteins are thermodynamically destabilized, a wide range of unrelated peptides and proteins have been observed to form artificial fibrillar materials *in vitro* that are characterized by a quaternary amyloid structure and this has led to the design of functional nanomaterials.

Amyloid materials, which are present in cells and in the extracellular space, represent a class of nanoscale structures that have various functional and pathological roles (Table 2; Fig. 2)<sup>15–17,19,21</sup>. As such, amyloid nanostructures are increasingly viewed as a general alternative form of protein structure that is different from, but in many cases no less organized than, the native states of proteins<sup>1,22</sup>. Moreover, this type of structure does not depend primarily on highly evolved side-chain interactions, but rather on universal physical and chemical characteristics that are inherent in the nature of all polypeptide molecules such as the propensity for hydrogen bonding in the backbone<sup>23</sup>.

Understanding why certain material characteristics have been conserved or refined over millions of years of evolution underlies many fundamental questions in biology, but this knowledge is also required to develop methods for using artificial proteinaceous nanostructures as practical functional materials. The question of material selection becomes particularly intriguing from a technological

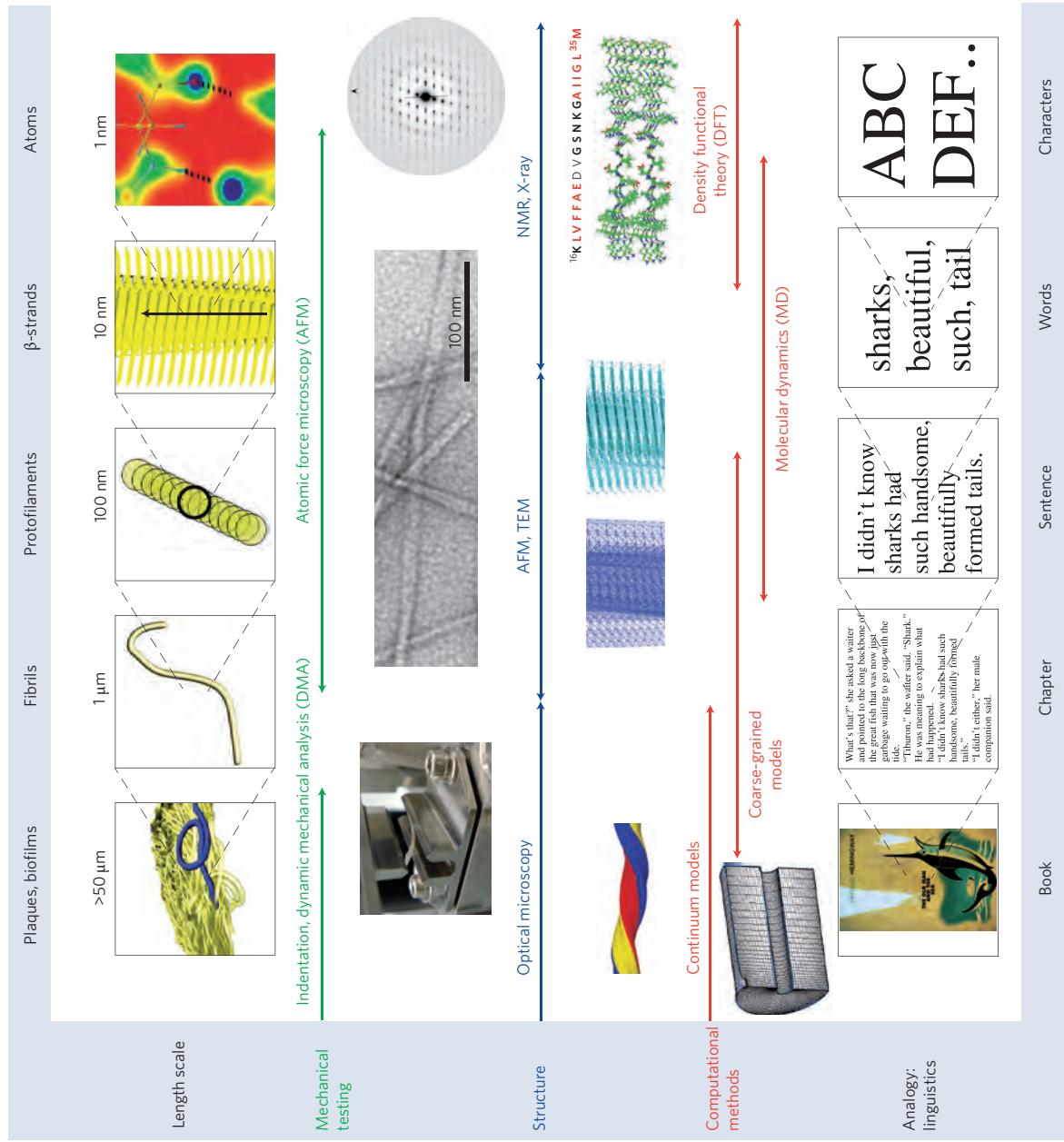
viewpoint because many protein materials are formed from scarce amounts of building blocks (that is, small volumes of material), from a few distinct building blocks (for example, only 20 natural amino acids) and from weak bonding (for example, hydrogen bonding), and are typically formed under severe energy constraints<sup>24</sup>. Because amyloid materials consist of generic assemblies of normally soluble proteins bound together in a simple periodic structure defined by main-chain hydrogen-bonding constraints, studying them should shed new light on the material selection criteria that shape more complex proteinaceous materials in nature.

## Nanomechanics of amyloid materials

Mechanical properties are critical to our understanding of how materials contribute to a biological or synthetic system. These properties include the strength of adhesive forces between fibrils or their surrounding, their rigidity, or the maximum stress that they can sustain without breaking. In common with many biological materials, amyloid materials have hierarchical structures (Fig. 1a) from the molecular to the macroscopic scale. Mechanical testing on the nanoscale allows the study of amyloid materials on the single fibril level and provides the ability to directly probe the forces that bind individual proteins together in such materials.

Nanomechanical testing can be performed by using atomic force microscopy (AFM) to carry out indentation measurements with high lateral resolution. Measurements of the contact stiffness of phenylalanine nanofibrils using this method showed that they are characterized by a Young's modulus ( $E$ ) of 19 GPa (ref. 25), implying a comparatively high stiffness (Fig. 3). This type of nanomechanical manipulation by AFM also allows the miniaturization of standard three-point bending testing. To this effect, the nanofibril structures are suspended over nanoscale gaps and an AFM tip is used to load the beam<sup>26,27</sup>. Experiments probing fibrils assembled from the protein insulin yield  $E = 3.3 \pm 0.4$  GPa (ref. 28). For nanofibrils with smaller diameters of only a few nanometres, it can be more challenging to probe their mechanical properties by directly applying a force and measuring resulting displacements because they are more fragile and flexible. Therefore, alternative methods to probe mechanical properties have been applied.

<sup>1</sup>Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK, <sup>2</sup>Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave, Room 1-25A &B, Cambridge, Massachusetts, USA, <sup>3</sup>Center for Computational Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts, USA, <sup>4</sup>Center for Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts, USA.  
\*e-mail: tpjk2@cam.ac.uk, mbuehler@MIT.EDU



**Figure 1 | The hierarchical structure of amyloid materials.** Upper panel: Five different levels of hierarchy in the structure of amyloid materials. Second and third panels: Different experimental and computational analysis tools for studying the mechanical properties and structure of amyloid materials on the different length scales. Second panel: The image on the left shows an artificial amyloid film being tested in a DNA setup. The transmission electron micrograph represents an Alzheimer's amyloid  $\beta$ -fibril. The image on the right shows the diffraction pattern from amyloid-like nanocrystals formed from the peptide GNNQQNY from the N-terminal region of the yeast prion protein Sup35. Figures in the second panel reproduced with permission from: left, ref. 93, © 2010 NPG; middle, ref. 14, © 2008 NPG; right, ref. 102, © 2003 Elsevier. Third panel: The image above 'Continuum models' represents a large-scale atomistic model of an Alzheimer's  $\beta$ -amyloid fibril; the image above 'Coarse-grained models' shows a coarse-grained elastic network model of an amyloid fibril<sup>32</sup>; the image above 'Molecular dynamics (MD)' depicts a molecular dynamics (MD) simulation of an Alzheimer's  $\beta$ -amyloid peptide oligomer (reproduced with permission from ref. 48, © 2002 NPG); and the image below 'Continuum models' shows a finite element model of an amyloid fibril (reproduced with permission from ref. 25, © 2005 ACS). Bottom panel: Hierarchical structure in linguistics as an analogy to demonstrate how functional properties emerge owing to the hierarchical assembly of simple building blocks. Figures in the bottom panel reproduced with permission from ref. 103, © 2007 Random House.

On the nanoscale, because elastic energy associated with small strain deformation becomes comparable to the thermal energy, spontaneous fluctuations of the fibril shape can occur and the statistical analysis of these geometric fluctuations can be used to measure mechanical properties<sup>29,30</sup>. Such an analysis has been carried out for different amyloid fibril systems and yields an  $E$  range of 0.2–14 GPa<sup>31,32</sup>. Further methods, including high-pressure X-ray diffraction studies<sup>33</sup> and statistical analysis of electron microscopy images of individual fibrils<sup>34</sup> show that amyloid fibrils commonly possess an elastic modulus in or close to the GPa range (Fig. 3).

This fluctuation analysis approach to nanomechanical characterization is advantageous over direct mechanical manipulation because fluctuations are evaluated on a length scale that greatly exceeds the tip radius; knowledge of the size or shape of the AFM tip is not required to interpret the measured data. However, the presence of a surface or of interactions between individual structures has the potential to lead to fluctuations other than those given by simple thermal motion of individual fibrils, and can represent an experimental challenge<sup>35</sup>.

Figure 3a summarizes the bending rigidity versus moment of inertia for covalent materials (the orange region), strong

**Table 1 | Summary of key terms related to amyloid materials science and nanotechnology.**

Term/concept	Definition	Relevance for amyloid and/or context, example	Nanotechnology relevance and/or potential applications
Cross- $\beta$ structure	Basic structural motif of amyloid fibrils consisting of $\beta$ -strands oriented perpendicular to the fibril axis.	Responsible for many of the nanoscale characteristics of amyloid fibrils, including their high elastic modulus.	Generic structural motif, mechanically and chemically stable owing to high density of H-bond clusters.
Amyloid fibril or fibre	Stack of cross- $\beta$ -motifs to form an elongated nanostructure. Diameters: typically 2–15 nm; length: typically 0.5–10 $\mu$ m.	Forms the basis for larger-scale assemblies, for example, amyloid plaques; see Fig. 1 or Fig. 5 for illustration.	Provides the basis for patterning constituents that do not form organized nanostructures on their own, such as metal particles or biochemically active agents (examples shown in Fig. 3b–h).
Amyloid plaques and biofilms	Micro to macroscale deposits of amyloid material in tissue, for example, the brain in Alzheimer's disease or in bacterial biofilms.	Assemblies of amyloid fibres; see Fig. 1 for several examples.	Control of patterning of amyloid fibrils in artificial films or on a surface enables the realization of different functional properties (Fig. 5a), for example, coatings, optical and electrical.
Prion	Proteinaceous particle responsible for the transmission of infectious conditions.	Prions propagate by transmitting a misfolded protein state (an amyloid structure in the case of yeast and probably a structurally related molecular species in the case of mammalian prions).	Design of self-replicating systems based on proteins and peptides.
Prionoid	General class of self-propagating protein elements but that lack microbiological transmissibility.	Many amyloid fibril systems have the ability to act as prionoids.	Self-regenerating and self-healing materials.
Strains	Classes of amyloid structures formed from polypeptide chains with an identical sequence but possessing a different three-dimensional packing.	Structural basis for the propagation of amyloid forms with different properties formed from a given protein.	Capacity to change functional properties of a material by assembling into different structures based on the same building blocks.
Mutability, tunability	Capacity to change functional properties of a material based on external signals (pH, light, for instance); in contrast to tunability, which is the capacity to change material reversibly during use.	Spontaneous growth of amyloid plaques in the context of disease states leading to changes in the properties of tissues.	Multifunctional materials with switchable states, for example, substrate patterning for tissue engineering, nanoscale valves and switches.
Optimality	Adaptation to reach a desired characteristic while respecting a set of restrictions.	Maximal stiffness of protein materials based on weak non-covalent bonding (performance limit), see Fig. 3.	Design of stiff materials based on peptide and protein constituents.
Evolvability	Ability to acquire new functions or features in response to changed conditions.	Self-replicating mechanisms associated with prion disease.	Design of self-adaptive materials, autonomous systems.
Universality	Occurrence of structure or mechanism (generally: protocol) in a great variety of systems.	Universal features in amyloid materials include cross- $\beta$ structure, hydrogen bonding, fibrillar nature and uniform diameter.	Universality of constituents underlies the ability to serve as generic building blocks.
Diversity	Occurrence of structure or mechanism in many guises, commonly linked to a particular functionality.	Diverse features in amyloid materials including sequence-dependent chemical nature of the side chains.	Use of diverse hierarchical structures enables the development of functional properties unique to a particular system (Fig. 5a) despite the presence of few universal building blocks (universality-diversity-paradigm).

non-covalent interactions (blue region) and weak non-covalent interactions (green region). The relationship of amyloid structures to these general material classes can be seen from the blue points that represent amyloid fibrils (different symbols denote data from different studies) and grey symbols are other materials<sup>25,31–34,36–38</sup> shown for comparison. In some cases where the  $E$  value of amyloid fibrils has been probed in the direction perpendicular to the fibril axis, significantly lower values have been reported (grey inverted triangles in Fig. 3a)<sup>39</sup>. This indicates a high level of anisotropy, and the lateral mechanical response is likely to be highly sensitive to the interactions between the component protofilaments rather than defined primarily through the intrinsic interactions within the  $\beta$ -sheet-rich amyloid core.

Measurements of the nanomechanics of less ordered forms of amyloid structures, known as protofibrillar assemblies, show<sup>35,38</sup> that they possess a lower value of  $E$  compared with mature amyloid

fibrils, indicating a different level of molecular organization in their core structure. Further detailed analysis of the shape fluctuations of protofibrillar assemblies indicate that heterogeneous populations can exist even within the same species<sup>38</sup>. This fact may be a manifestation of the general tendency of amyloid materials towards polymorphism through the existence of many strains of fibrils formed by the same polypeptide sequence but characterized by subtle changes in the molecular packing of the chains within the fibrils<sup>4,40,41</sup>. Even small changes in the conditions during fibril growth can bias the system towards the formation of a different strain. Examples include quiescent versus agitated incubation conditions<sup>38</sup>, the presence of cosolvents<sup>42</sup> and temperature<sup>41</sup>. The structural features that are characteristic of a specific strain are in many cases transmitted to a new generation of fibrils that are grown when a monomeric solution of the protein is exposed to preformed seed fibrils of a given strain. Such strains can in many cases distinguished through

**Table 2 | Examples of different forms of amyloid material (adapted from ref. 70).**

Context/function	Protein	Primary localization or organism
Functional amyloid		
Control of transcription	Ure2p	Yeast
Control of translational termination	Sup35	Yeast
Biofilm production	Curlin	<i>E. coli</i> bacteria
Heterokaryon incompatibility	Het-S	Fungi
Pituitary secretory granules	Secretory hormones	Pituitary gland
Melanin biosynthesis	Pmel17	Melanosomes in mammalian skin
Pathological amyloid		
Alzheimer's disease	Amyloid- $\beta$ ( $\alpha\beta$ ) peptide	Extracellular (central nervous system (CNS))
Parkinson's disease	$\alpha$ -synuclein	Intracellular (CNS)
Tauopathies, Alzheimer's disease	Tau protein	Various tissues, including liver, heart, spleen, and kidney
Systemic amyloidosis	Lysozyme, transthyretin, serum amyloid A, immunoglobulin light chain and others	Various tissues, joints
Haemodialysis-associated amyloidosis	$\beta2$ -microglobulin	Extracellular (CNS)
Prion diseases	Prion protein	Intracellular (CNS)
Huntington's disease	Polyglutamine-rich proteins	Extracellular (pancreas)
Type II diabetes	Amylin	

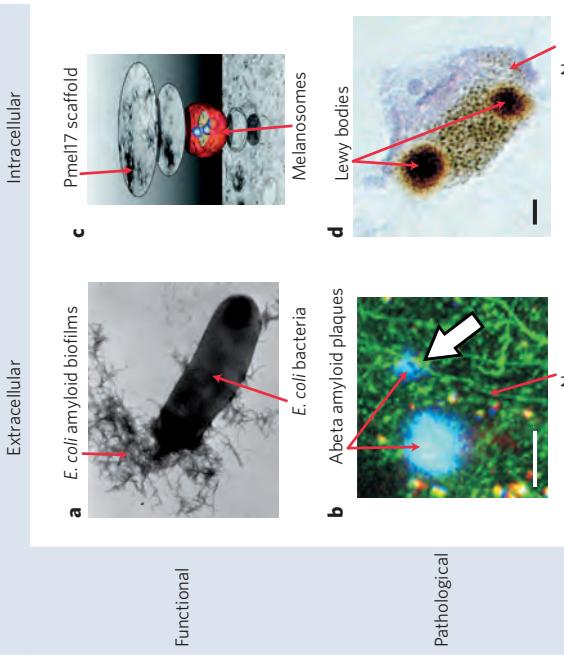
their mechanical properties, but a remarkable discriminatory power between even subtle changes in structure in amyloid materials can be achieved through the use of recently developed conjugated polyelectrolytes<sup>43</sup>, which possess conformation-dependent fluorescence spectra when bound to amyloid fibrils.

General trends that underlie the modulation of the material properties of amyloid fibrils can be identified. For instance, the highest  $E$

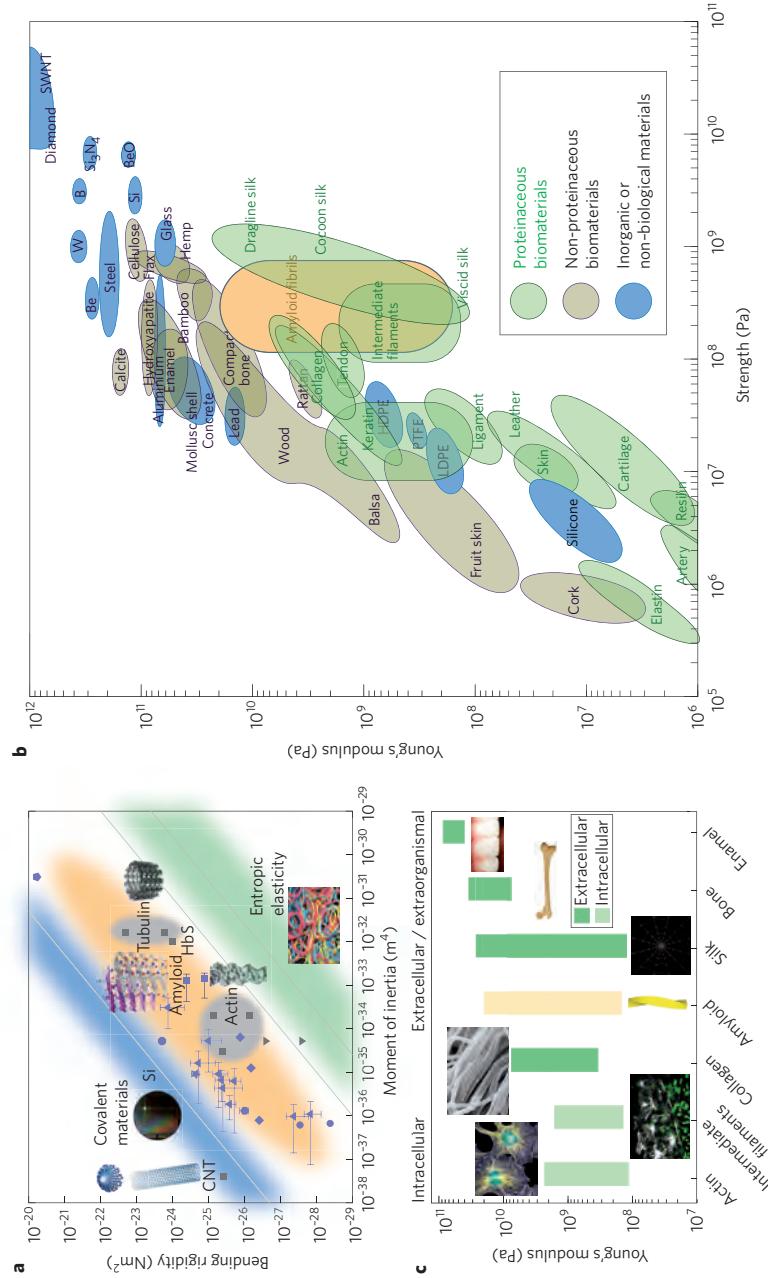
values are observed for short peptides such as diphenylalanine (two residues)<sup>25</sup>, yeast prion fragment (seven residues) and transthyretin fragment (eleven residues)<sup>32</sup>, whereas the lower values tend to be observed for fibrils from longer sequences such as  $\alpha$ -lactoglobulin,  $\beta$ -lactalbumin<sup>32</sup> and 'HypF<sup>38</sup>' (Fig. 3). These observations point towards the role of increasing structural disorder arising from the constraints accompanying the packing of increasingly long polypeptides into fibrillar structures; where such effects lead to a less effective search for strong intermolecular bonding, a lower modulus can result. Similar observations have been made on larger scales, where computation studies show that highly organized assemblies of short amyloid fibrils have a larger modulus than less organized superstructures composed of longer amyloid fibrils<sup>44</sup>.

The mechanical failure of amyloid materials is at present less well characterized than their linear and elastic deformation at small strain. Experimental values have been obtained from destructive mechanical testing of fibrils using AFM and from the observation of filament fragmentation owing to solvent-imposed shear fields. These measurements yield estimates for the tensile strength of insulin amyloid filaments in the range of 0.1–1 GPa<sup>28,45</sup>.

*In silico* models of amyloid materials have developed into powerful tools to complement experimental methods from first principles. Suitable computational methods include quantum mechanical<sup>46</sup>, molecular dynamics<sup>11,13,36,47–49</sup>, coarse-grained<sup>44,50–53</sup> and continuum approaches<sup>25,52</sup>. A key focus of computational modelling has been the identification of molecular details of the initial stages of amyloid aggregation and defining the nanomechanics of amyloid materials. *In silico* studies of amyloid nanomechanics indicate that the values of  $E$  are high for non-covalent materials, in the range of 10–20 GPa, and originate largely from the dense backbone hydrogen-bonding network that extends throughout the core of the structures<sup>35,37,54</sup>. Interactions between groups in the side chains of the polypeptide molecule can further modulate the mechanical properties of the resulting fibrils. This is particularly apparent in cases where side chains can significantly contribute to the hydrogen-bonding patterns, such as in the case of yeast prion fragments<sup>35</sup>. In addition to interactions within the sheets, side-chain interactions play a crucial role in the interactions between the sheets; these interactions typically take the form of a steric zipper, where the complementarity of the adjacent interfaces drives their lateral association<sup>4</sup>.



**Figure 2 | Classification of amyloid materials.** Amyloid materials can be extracellular (**a, b**) or intracellular (**c, d**), and functional (**a, c**) or pathological (**b, d**). **a**, Functional amyloid<sup>70</sup> in biofilms produced by bacterial species such as *E. coli* and certain *Salmonella* spp. **b**, Amyloid plaques as seen in a mouse model of Alzheimer's disease (scale bar, 20  $\mu$ m). The large white arrow shows a newly formed plaque. **c**, Transmission electron microscope image of Pmel17 scaffolds in melanosomes involved in the biosynthesis of melanin. **d**, Lewy bodies, pathological protein aggregates that develop in neurons in Parkinson's disease (scale bar, 8  $\mu$ m). Figures reproduced with permission from: **a**, ref. 79, © 2007 NAS; **b**, ref. 77, © 2008 NPG; **c**, ref. 20, © 2009 ASEM; **d**, ref. 21, © 1997 NPG.



**Figure 3 | Mechanical properties of amyloid fibrils in comparison to biological and inorganic or non-biological materials.** **a**, Bending rigidity versus moment of inertia for covalent materials (blue region), strong non-covalent interactions (such as hydrogen bonds, orange region) and weak non-covalent interactions (green region). Blue points are amyloid fibrils (different symbols denote data from different studies) and grey symbols are other materials<sup>25,31–34,36–38</sup>. Grey inverted triangles show the values for the response measured in the perpendicular direction to the fibril axis<sup>39</sup>. The upward triangles show data for one- and two-filament forms of bovine insulin, B-chain of bovine insulin, hen-egg-white lysozyme, bovine  $\beta$ -lactoglobulin, Alzheimer's amyloid  $\beta$ -peptide residues 1–42, GNNQQNY fragment of the yeast prion sup35, and human transthyretin residues 105–115 (all experimental)<sup>32</sup>. The pentagons show data for diphenylalanine (experimental)<sup>25</sup>, octagons for insulin (experimental)<sup>33</sup>, hexagons for ac-[RARA]DADA2-arm self-assembling peptide (simulation)<sup>36</sup>, stars for Alzheimer's amyloid  $\beta$ -peptide residues 1–40 (simulation)<sup>37</sup>, and lozenges show data for Alzheimer's amyloid  $\beta$ -peptide residues 1–40 (experimental)<sup>34</sup>. Figures reproduced with permission from: amyloid, ref. 104, © 2005 NPG; actin, ref. 105, © 2008 NPG; tubulin, ref. 106, © 2002 Elsevier; images courtesy of silicon wafer © istockphoto.com/photomick; elastic bands © istockphoto.com/shank/-ali. **b**,  $E$  (which corresponds to stiffness) versus strength<sup>56,57</sup> for a range of different materials. Covalent and metallic bonding results in the stiffest and strongest materials, with diamond and single-wall carbon nanotubes (SWNTs) being the best performers. Silks are the strongest and stiffest protein materials, followed by amyloid and collagen; and significantly more-rigid materials (for example, bone) contain minerals. Amyloid fibrils are shown in orange to distinguish them from other materials. **c**, Range of values of  $E$  for seven different classes of biological materials inside and outside the cell. The stiffest materials (such as collagen, bone, enamel and silk) are found outside the cell. Images courtesy of: actin © NIGMS/Torsten Wittmann; intermediate filaments © NIGMS/Evan Zamir; collagen © fei.com/Paul Gunning; silk © istockphoto.com/blackjacket3d; bone © istockphoto.com/dwwithers; enamel © istockphoto.com/shironosov.

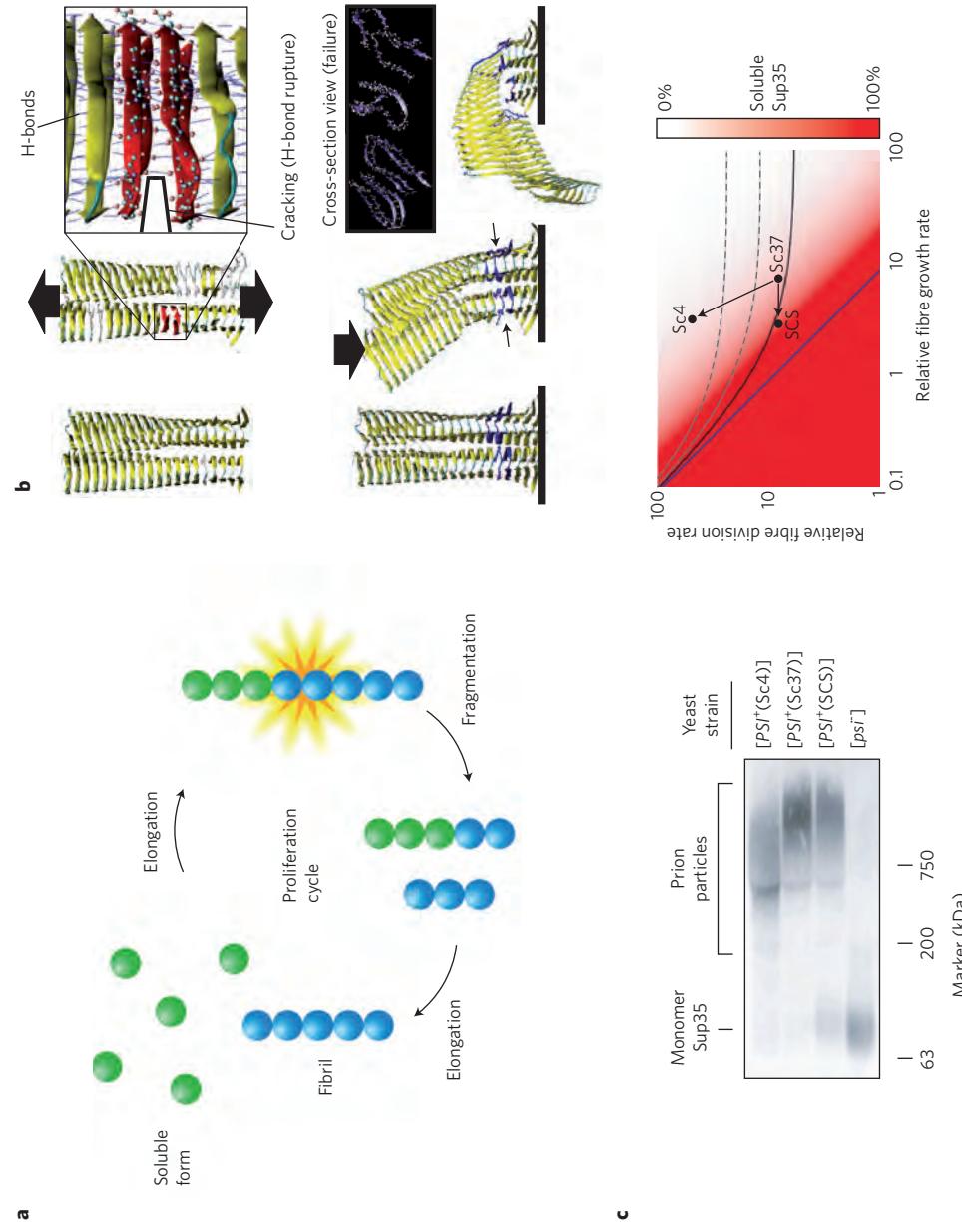
### Relation to other biological materials

The mechanical properties of materials intimately reflect the nature of the fundamental intermolecular interactions that bind their constituents into larger-scale hierarchical systems. In amyloid materials this connection offers general insights into the characteristics of interactions in natural or artificial proteinaceous materials. Amyloid fibrils can be seen to form a reference class of structures where the core of fibrils consists of arrays of  $\beta$ -strands that involve essentially every residue in the core domain, resulting in continuous  $\beta$ -sheets that can extend over thousands of molecular units, a unique situation in biological materials.

Many amyloid fibrils, especially those formed from proteins that do not undergo assembly towards the amyloidogenic state in nature, have not been directly influenced by evolutionary pressures. As such, these proteins give access to the inherent mechanical properties of  $\beta$ -sheets on supramolecular length scales without the influence of evolutionary adaptation that typically pervades the properties of natural protein materials. Compared with other

biological materials (Fig. 3)<sup>55–57</sup>, amyloid fibrils are remarkably stiff and possess values of  $E$  comparable to the most rigid proteinaceous materials in nature. Dragline silk, for example, has a value of  $E$  of up to 10 GPa, and collagen and keratin both exhibit  $E$  values approaching this value (Fig. 3b)<sup>45,58–60</sup>. Dragline silk has similarities from a structural point of view with amyloid fibrils, and the crystalline regions of silk<sup>61,62</sup>, which contribute to its high  $E$  value, consist of densely packed hydrogen-bonded  $\beta$ -strands analogous in many ways to those found in the core of amyloid materials.

Considerations based on the maximal density of intermolecular hydrogen bonds that can be achieved in proteins yield a limit for the material performance of proteinaceous structures in terms of  $E$  on the order of 10–20 GPa<sup>24,32</sup>. In agreement with this concept, materials that possess  $E$  values significantly above this limit contain covalent or metallic interactions with a significantly higher energy density than hydrogen bonds or similarly weak non-covalent interactions<sup>55</sup>. Nature has therefore, through the action of evolutionary processes, been able to optimize proteinaceous materials to be close to



**Figure 4 | Fragmentation, aggregation and the kinetics of amyloid growth.** **a**, Schematic illustrating the formation of amyloid fibrils from soluble monomers. On fragmentation, newly formed ends of amyloid fragments serve as seeds for further elongation, and the process repeats<sup>63</sup>. **b**, Molecular mechanism of fragmentation of the cross- $\beta$  core of amyloid fibrils<sup>65</sup>. Upper part: failure owing to tension, through the opening of a crack due to the breaking of hydrogen bonds. Lower part: failure under compression owing to fluid shear forces or thermal vibrations that excite bending or stretching modes. **c**, Studies of the propagation of three strains of yeast prion (Sc4, Sc37 and SCS) show that there exists an intimate connection between mechanics and the growth kinetics of prion aggregates. These amyloid fibrils formed from the protein Sup35 propagate most effectively in yeast cells in cases where their intrinsic fragmentation rate was high and the mechanical strength correspondingly low. The gel electrophoretic analysis of prion particle size (**c**, left) shows that the greater fragility of Sc4 fibrils in a smaller size distribution of the aggregates, but more effective overall growth and sequestering of the monomer relative to the less fragile SCS fibrils. The [*psi*<sup>-</sup>] case represents a control system with no aggregates. The right panel shows the differences in the fibril breakage rates and growth rates. Reproduced from ref. 41, © 2006 NPG. **d**, The tensile strength in GPa (red numbers, results from computational modelling) of amyloid fibrils of various lengths (black numbers): it can be seen that longer fibrils are weaker<sup>64</sup>. The plot also summarizes characteristic length scales of amyloid fibrils, including the layer spacing (distance between  $\beta$ -sheet layers in the core), the helical pitch in the twisted geometry of the fibrils, and the persistence length (data shown here are for Alzheimer's  $\beta$ -amyloid fibril). As the length of the amyloid fibrils that assemble into plaques is varied, the structure of the assembled plaques changes from being highly organized and regular (green, right) to more entangled and randomized (red, left)<sup>4,28,35,44,46</sup>. This is because of the competition between bending and interfibril adhesion; in shorter fibrils the interfibril adhesion forces are strong enough to align fibrils in parallel, whereas for longer fibrils the entangled arrangement is stable.

the theoretical performance limits. It is apparent, however, that the rigidity of cytoskeletal filaments<sup>57–64</sup> does not extend up to the maximal values achievable for biological materials (Fig. 3c). Intracellular materials are subject to an additional requirement that assembly must be readily reversible under physiological conditions to ensure cellular motility and compatibility with the intracellular environment. Indeed, the strong intermolecular bonding in the materials that exhibit the highest moduli — cellulose, bone, keratin, silk — seems to be incompatible with rapid and dynamic disassembly characteristic of functional intracellular structures, and these materials are primarily found extracellularly and in roles where controlled rapid breakdown is not required.

### Failure is the key to growth

The interplay between nanomechanics of amyloid fibrils and their growth kinetics is increasingly appearing as a key factor for the overall conversion of proteins from their soluble to fibrillar states in many amyloid-related phenomena, including the transmission of prions<sup>41,63</sup>. This surprising connection emerges through the importance of filament fragmentation as one of the fundamental factors controlling the proliferation of amyloid fibrils. Because the growth of such structures occurs by addition of soluble proteins to fibril ends, the number of free ends effectively governs the overall conversion reaction (Fig. 4). In agreement with this concept, even small shear forces, which favour fragmentation of filaments but are likely to leave other system parameters largely invariant, are generally observed to significantly accelerate amyloid growth. Despite the substantial strength of amyloid structures, as their length increases through growth, the possibility of structures fragmenting either spontaneously through thermal fluctuation or through the action of cellular agents becomes increasingly likely<sup>64</sup>. Indeed, *in silico* studies of amyloid fibril failure suggested that longer fibrils are more brittle elements that show a greater propensity towards failure as their length increases (Fig. 4d)<sup>24,65</sup>.

The role of fragmentation is particularly marked for prion-type aggregates. These aggregates are characterized by their remarkable ability to proliferate and infect organisms where the soluble, non-aggregated form of the prion protein is available<sup>66</sup>. Examples of such behaviour are known in a variety of organisms including fungi and mammals. For this class of structure, the spontaneous generation of new fibrils through primary nucleation is generally very slow, and secondary pathways such as fragmentation are therefore critical in generating new growth sites as the reaction progresses (Fig. 4c)<sup>63</sup>. Amyloid fibrils formed from the yeast protein Sup35 were shown to propagate as prions most effectively in yeast cells in cases where their intrinsic fragmentation rate was high and the mechanical strength correspondingly low<sup>41,65</sup>. Although mammalian prion systems are more complex than their yeast analogues, trends indicative of a similar connection have been identified. Indeed, the time before disease development in a mouse model of prion disease was found to be inversely correlated with the aggregate stability. This suggests a lower stability contributing to a higher fragmentation rate, more effective prion proliferation and thus a more rapid onset of disease symptoms<sup>67</sup>. Interestingly, recent evidence points towards remarkable underlying similarities between prion conditions and other amyloid disorders such as Alzheimer's disease<sup>68,69</sup>. Indeed, although under most circumstances true infection between organisms remains confined to prion type aggregates, it has been demonstrated that amyloid fibrils formed from many other proteins can also be released from cells where they have grown and can infect neighbouring cells, giving rise to the concept of a prionoid as a general self-propagating protein element<sup>66,70</sup>. The role of material failure therefore, shown to be crucial in the proliferation of prions as discussed above, may also be important in the development of other amyloid pathologies. Furthermore, many recent studies suggest that the deleterious activity associated with protein aggregates stems from aggregates of small molecular weight. Although their source has still to be established

with certainty, it could be hypothesized that their release from larger plaques to form a halo of small oligomers<sup>71</sup> through fragmentation may be a significant contribution<sup>72</sup>.

### Natural uses of amyloid as a functional material

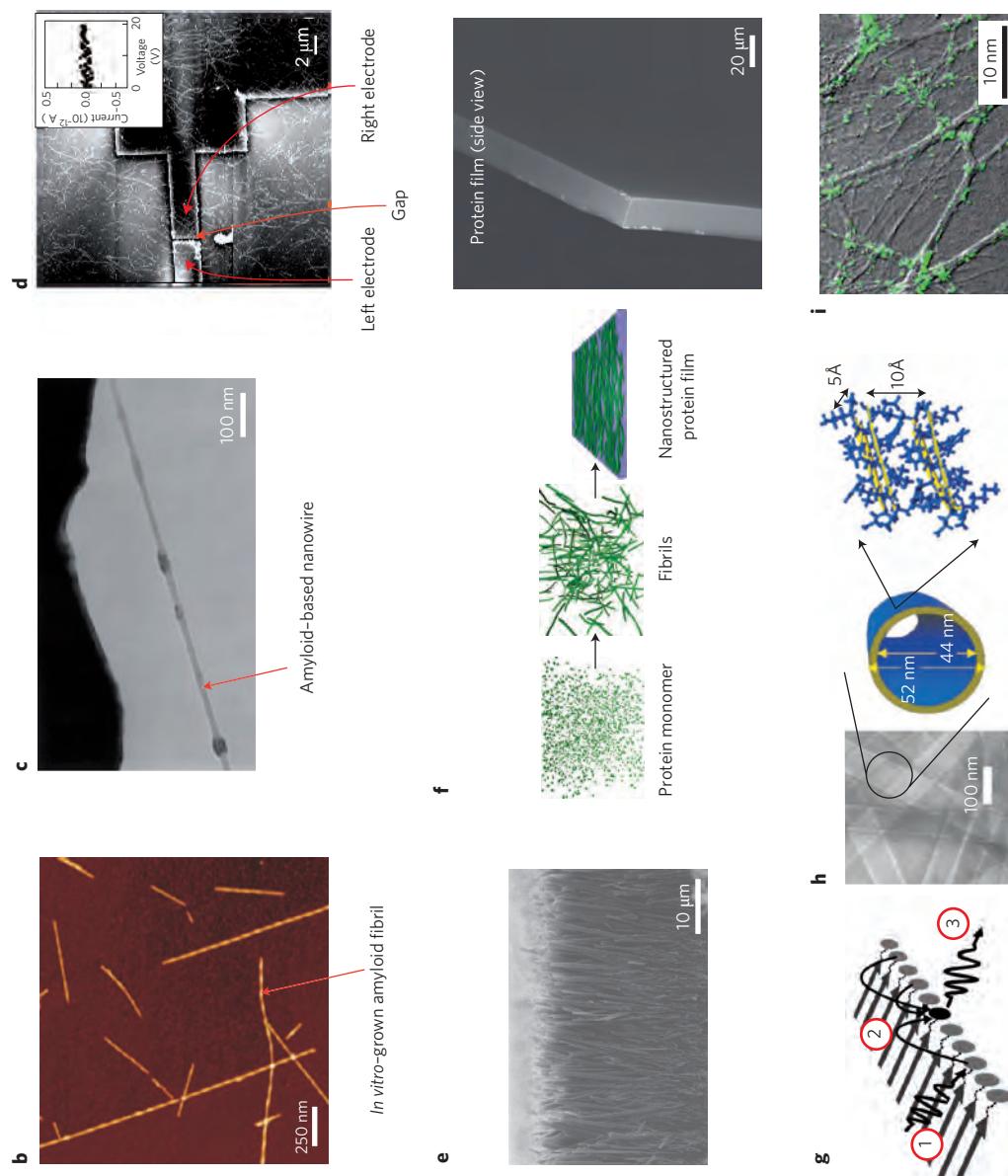
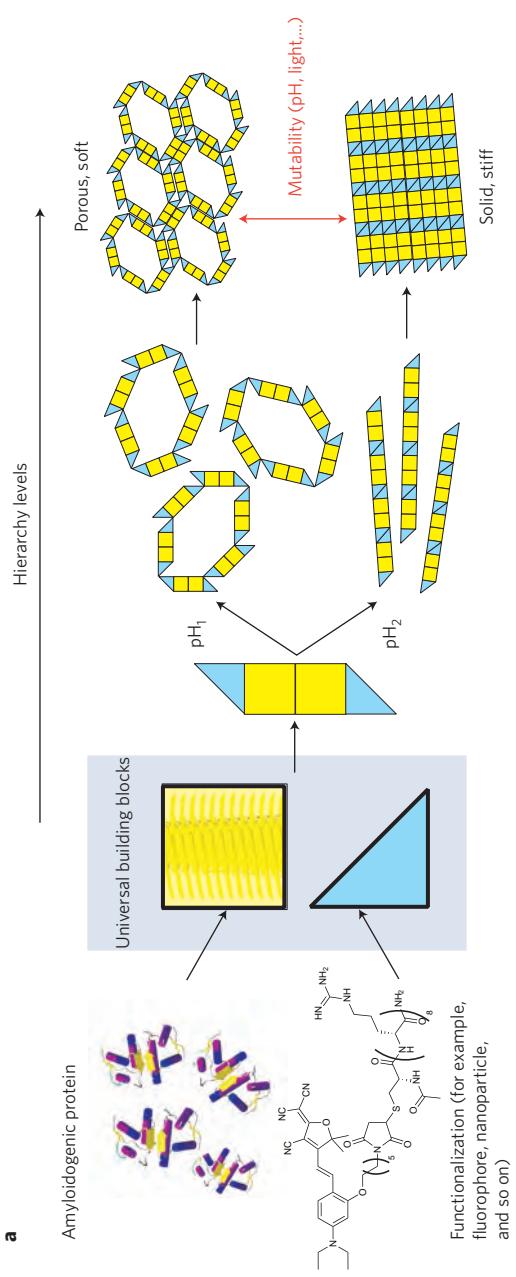
Ever since their discovery, amyloid fibrils have been primarily associated with pathological behaviour<sup>10,12</sup>, but amyloid structures are in fact also naturally found in many functional roles<sup>5,16,18,19,73,74</sup>. Prominent examples of such functional amyloid materials include bacterial coatings<sup>75</sup>, catalytic scaffolds<sup>16</sup>, agents mediating epigenetic-information storage and transfer<sup>76</sup>, adhesives<sup>19</sup> and structures for the storage of peptide hormones<sup>74</sup>. The existence of functional amyloid is highly significant, because it demonstrates that the amyloid fold is not unequivocally connected to toxicity and disease<sup>8,10,77</sup>.

Rather, it represents a more general ordered state for proteins that is often energetically favourable compared with the soluble native state<sup>1</sup>. These discoveries are remarkable from a biological point of view as they highlight the delicate distinction between pathology and physiology, but they are also intriguing because they point to the possibility of using amyloid as a powerful platform for the design of new nanomaterials.

Functional amyloid was initially identified in single-cell organisms. Many such applications are characterized by the use of amyloid as a structural material. Bacteria, including *Escherichia coli* (*E. coli*), use an extracellular amyloid material, curli fibres, as the basis for generating a matrix for surface adhesion and interactions with other bacteria (Fig. 2a)<sup>75,78,79</sup>. Several species of fungi, including the common baker's yeast, use amyloid fibrils for storing epigenetic information. Here amyloid structures implement a switch, the state of which is defined through the presence of an amyloidogenic protein either in its soluble or fibrillar form. Aggregates in the cytoplasm will be transmitted to the daughter cells during cell division; when aggregates are present they will act as seeds to promote the conversion of the relevant protein into its fibrillar state also in the daughter cells, and conversely when no aggregates are present in the mother cell, the amyloid conversion is significantly less likely. The information encoded directly through the bistability in the configuration of a protein is therefore inheritable. Another natural use of amyloid as a functional material is as a component of extracellular polymeric substances secreted by algae that mediate surface adhesion<sup>19</sup>.

Functional amyloid is not a phenomenon restricted to single-cell organisms. Indeed, an example of amyloid scaffolds possessing a key physiological role has emerged in the biosynthesis of melanin in humans<sup>16</sup>. Melanin synthesis in melanosomes, specialized secretory organelles in melanocytes in the skin and eyes, occurs through polymerization of small molecule precursors. This process is catalysed by a scaffold of amyloid fibrils formed from the protein Pmel17 (refs 18,20; Fig. 2c). The role of the amyloid scaffold is thought to be related to a templating process, where the effective concentration of precursor molecules is increased through their adsorption onto the catalytical scaffold. The efficiency of the reaction may be enhanced through the orientational preference provided by amyloid fibrils. Amyloid fibrils may have many other functional roles in mammalian systems, and recently a further natural application was revealed in the form of high-density packing of peptide hormones in the secretory granules in pituitary glands<sup>74</sup>. Here the formation of amyloid fibrils helps to isolate peptides with a given sequence as well as stabilize the hormones during storage before secretion.

The amyloid fold can provide an effective basis for the formation of robust fibrillar materials through self-assembly. A key open question is the lack of toxicity to the host organism synthesizing functional amyloid when compared with the deleterious effects associated with the deposition of amyloid in disease states. Some insights that could resolve this paradox come from the realization that the formation of functional amyloid is frequently under tight



**Figure 5 | Examples of functional synthetic amyloid materials.** **a**, Amyloid formation from a native protein results in universal building blocks that can be assembled and functionalized (for example, with fluorophores or metal particles) into larger and more diverse structures. Different structures can be achieved through changes in pH or other processing conditions. **b**, AFM image of synthetic amyloid fibrils generated *in vitro*<sup>35</sup>. **c–f**, The fibrils can assemble to form conducting nanowires (**c**,<sup>36</sup> surface coatings (**e**,<sup>31</sup>) and nanostructured protein films (**f**)<sup>33</sup>. **g**, Schematic showing a light-harvesting nanostructure generated from amyloid fibrils<sup>35</sup>. Light harvesting occurs by means of absorption of a photon by the donor (1), followed by non-emissive transfer to an acceptor through resonance energy transfer (2). The energy is released by the acceptor as a photon (3). **h**, Hollow nanotubes made from amyloid structures could be used to develop new nanoscale antennas<sup>36</sup>. **i**, Optical image of active neuron synapses (labelled with a green fluorescent lipophilic probe) whose growth was stimulated by  $\beta$ -sheet-rich scaffolds<sup>38</sup>. Figure reproduced with permission from: **b**, ref. 35, © 2007 AAAS; **c**, ref. 86, © 2006 ACS; **d**, ref. 88, © 2003 NAS; **e**, ref. 91, © 2009 NPG; **f**, ref. 93, © 2010 NPG; **g**, ref. 95, © 2009 ACS; **h**, ref. 96, © 2008 RSC; **i**, ref. 98, © 2000 NAS.

control and takes place under conditions that favour rapid and effective polymerization<sup>75</sup> such that the formation of potentially toxic low-molecular-weight oligomers or other intermediates is avoided.

### Synthetic amyloid-based nanostructures and applications

Multifunctionality in natural proteinaceous materials is usually produced by assembling a few simple (and often lower-performance) elements into structures that are characterized by multiple length-scales (Fig. 5a). This is completely different from the traditional engineering approach, which involves the use of a large number of distinct (and often high-quality) building blocks. An intriguing aspect of amyloid mechanics is the manner by which weak non-covalent intermolecular chemical bonds generate mechanically strong materials. *In silico* studies show that the key to the creation of strength from weak hydrogen bonding is their grouping into clusters, arranged in  $\beta$ -strands with a relatively short length. This universal design feature allows hydrogenbonds to work cooperatively and reach maximum strength<sup>24,83</sup>. Many of the natural applications of amyloid materials discussed above capitalize on their robust properties and readily self-assembling nature in the absence of external energy input such as ATP; these characteristics are also of interest for synthetic biomaterials, and the discovery of functional amyloid has provided the inspiration for the development of artificial amyloid materials.

An advantage of biological self-assembly is the ability to generate and control structure on the nanoscale, as shown schematically in Fig. 5a. This natural propensity for nanoscale organization into fibrils (Fig. 5b)<sup>35,84</sup> can be used to template other materials that do not on their own possess a propensity to form ordered structures on that scale, such as metal particles<sup>85</sup>. This strategy, which arranges universal building blocks into hierarchical structures to create diverse functional materials that are similar to natural materials, offers opportunities in hierarchical *de novo* material design. The power of this principle has been demonstrated in the fabrication of conductive nanowires, where the self-assembling peptide fibrils act as templates for the deposition of metals on the outside of structures to yield electrically conducting wires (Fig. 5c,d)<sup>86–88</sup>. Diphenylalanine nanotubes are of particular interest in this context because these structures contain a hollow core; deposition of metal both within and on the outside of these structures results in coaxial nanowires with electromagnetic properties<sup>89</sup>. The assembly of amyloid fibrils into larger-scale structures also provides the opportunity to realize new surface (Fig. 5e)<sup>90–92</sup> and bulk (Fig. 5f)<sup>93</sup> properties. The propensity of proteins to undergo multilevel hierarchical assembly opens up the possibility of larger-scale structures to be generated through self-assembly, while maintaining the accurate control of nanoscale organization<sup>93</sup>.

Another use of amyloid scaffolds has been demonstrated in the context of organic photovoltaics<sup>94</sup>. A challenge in the fabrication of such materials stems from the requirement to generate and control a large interfacial area between electron donor and acceptor materials where photocharges are created. Improved characteristics were reported for organic solar cells where amyloid fibrils were used as a template to orient the donor and acceptor polymers and to enhance the area of the donor–acceptor interface. The assembly of a peptide scaffold can also be used to drive the organization of host species attached to the peptides before assembly. This has enabled the creation of linear nanoscale arrays of fluorescent species that on illumination allow energy migration along the scaffold in the form of eximers. When binary structures are created that include both acceptor and donor groups in the same fibril scaffold, excitation of donor species by incident light allows energy transfer to acceptor sites where the energy can be converted back to light, and emission can be observed. Such structures can operate as light-harvesting materials (Fig. 5g)<sup>95</sup>. A recent study demonstrated the assembly of strong chromophores (colouring pigments) across

a paracrystalline amyloid network, which allows for precise ordering along the inner and outer compartment walls of an amyloid-based protein nanotube in a nanoscale antenna<sup>96</sup> (Fig. 5h).

The inherently fibrillar nature of amyloid materials alone can lead to useful modifications to the chemical availability of biologically active species. In particular, amyloid has been explored for controlled-release drug delivery, where longer-lasting action can be achieved through the slow dissociation of peptide nanostructures after administration. This principle has been demonstrated in the context of cancer therapy, based on, for example, gonadotropin-releasing hormone, the action of which can be significantly prolonged when fibrillar rather than monomeric material is administered<sup>97</sup>. The fibrillar structures of amyloid materials are also applied in three-dimensional tissue-culture scaffold design<sup>98</sup> (Fig. 5i), for brain repair through axon regeneration<sup>99</sup>, or for the modulation of cell adhesion<sup>100</sup>.

The paradigm of using hierarchical structures to create diversity of function out of simple, universal elements, as seen in amyloid nanomaterials, is in many ways analogous to linguistics (Fig. 1)<sup>101</sup>; a set of letters of the alphabet are used to form words, which are assembled into sentences, which are assembled into paragraphs, chapters and eventually an entire book. The construction of language exemplifies how the interplay of diversity and universality provide a powerful bottom-up design approach that can be applied to the development of nanomaterials (see Fig. 5).

### Conclusion

The nanoscale mechanics of amyloid fibrils underlies many aspects of the behaviour of these materials in their different biological contexts, including molecular medicine, biomechanics and functional bionanomaterials. Understanding the role of nanoscale material failure in the fragmentation of aggregates and the proliferation of amyloid fibrils in prion and related disorders offers insights into the structural and mechanistic differences between physiologically useful and pathological amyloid structures. The principle of hierarchical assembly in the amyloid scaffold can potentially be exploited to generate new forms of functional biomaterials with a wide range of physical properties in which a small number of abundant natural building blocks assemble to yield high-performance but environmentally benign materials.

### References

- Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24**, 329–332 (1999).
- Jaroniec, C. P. *et al.* High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **101**, 7111–7116 (2004).
- Luhrs, T. *et al.* 3D structure of Alzheimer's amyloid- $\beta$ (1–42) fibrils. *Proc. Natl. Acad. Sci. USA* **102**, 17342–17347 (2005).
- Sawaya, M. R. *et al.* Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
- Wasmser, C. *et al.* Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526 (2008).
- Sipe, J. D. & Cohen, A. S. Review: History of the amyloid fibril. *J. Struct. Biol.* **130**, 88–98 (2000).
- Tan, S. Y. & Pepys, M. B. Amyloidosis. *Histopathology* **25**, 403–414 (1994).
- Selkoe, D. J. Alzheimer's disease: Genes, proteins, and therapy. *Physiol. Rev.* **81**, 741–766 (2001).
- Antzutkin, O. N. *et al.* Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils. *Proc. Natl. Acad. Sci. USA* **97**, 13045–13050 (2000).
- Bucciantini, M. *et al.* Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).
- Tsai, H. H. *et al.* Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: significant role of Asn ladder. *Proc. Natl. Acad. Sci. USA* **102**, 8174–8179 (2005).
- Pepys, M. B. Amyloidosis. *Ann. Rev. Med.* **57**, 223–241 (2006).

13. Zanuy, D., Gunasekaran, K., Lesk, A. M. & Nussinov, R. Computational study of the fibril organization of polyglutamine repeats reveals a common motif identified in beta-helices. *J. Mol. Biol.* **358**, 330–45 (2006).

14. Paravastu, A. K., Leipman, R. D., Yau, W. M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proc. Natl. Acad. Sci. USA* **105**, 18349–18354 (2008).

15. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–66 (2006).

Overview of the principles of protein folding and misfolding.

16. Fowler, D. M. *et al.* Functional amyloid formation within mammalian tissue. *PLoS Biol.* **4**, e6 (2006).

Functional amyloid in the synthesis of melanin.

17. Fowler, D. M., Koulov, A. V., Balch, W. E. & Kelly, J. W. Functional amyloid from bacteria to humans. *Trends Biochem. Sci.* **32**, 217–24 (2007).

18. Kelly, J. W. & Balch, W. E. Amyloid as a natural product. *J. Cell Biol.* **161**, 461–462 (2003).

19. Mostaert, A. S., Higgins, M. J., Fukuma, T., Rindi, F. & Jarvis, S. P. Nanoscale mechanical characterisation of amyloid fibrils discovered in a natural adhesive. *J. Biol. Phys.* **32**, 393–401 (2006).

20. Watt, B. *et al.* N-terminal domains elicit formation of functional Pmel17 amyloid fibrils. *J. Biol. Chem.* **284**, 35543–35555 (2009).

21. Spillantini, M. G. *et al.*  $\alpha$ -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).

22. Fandrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165–166 (2001).

23. Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–90 (2003).

24. Keten, S., Xu, Z., Ihle, B. & Buehler, M. J. Nanoconfinement controls stiffness, strength and mechanical toughness of beta-sheet crystals in silk. *Nature Mater.* **9**, 359–367 (2010).

25. Kol, N. *et al.* Self-assembled peptide nanotubes are uniquely rigid bioinspired supramolecular structures. *Nano Lett.* **5**, 1343–1346 (2005).

26. Salvatet, J. P. *et al.* Elastic and shear moduli of single-walled carbon nanotube ropes. *Phys. Rev. Lett.* **82**, 944–947 (1999).

27. Kis, A. *et al.* Nanomechanics of microtubules. *Phys. Rev. Lett.* **89**, 248101 (2002).

28. Smith, J. E., Knowles, T. P. J., Dobson, C. M., MacPhee, C. E. & Welland, M. E. Characterization of the nanoscale properties of individual amyloid fibrils. *Proc. Natl. Acad. Sci. USA* **103**, 15806–15811 (2006).

29. Gittes, F., Mickey, B., Nettleton, J. & Howard, J. Flexural rigidity of microtubules and actin-filaments measured from thermal fluctuations in shape. *J. Cell Biol.* **120**, 923–934 (1993).

30. Wang, J. C. *et al.* Micromechanics of isolated sickle cell hemoglobin fibers: Bending moduli and persistence lengths. *J. Mol. Biol.* **315**, 601–612 (2002).

31. Adamcik, J. *et al.* Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. *Nature Nanotech.* **5**, 423–428 (2010).

32. Knowles, T. P. *et al.* An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533–1537 (2009).

33. Meersman, F., Cabrera, R. Q., McMillan, P. F. & Dmitriev, V. Compressibility of insulin amyloid fibrils determined by X-ray diffraction in a diamond anvil cell. *High Pressure Res.* **29**, 665–670 (2009).

34. Sachse, C., Grigorieff, N. & Fandrich, M. Nanoscale flexibility parameters of Alzheimer amyloid fibrils determined by electron cryo-microscopy. *Angew. Chem. Int. Ed.* **49**, 1321–1323 (2010).

35. Knowles, T. P. *et al.* Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **318**, 1900–1903 (2007).

Determination of the rigidities of nanofibrils formed from a wide range of peptides and proteins.

36. Park, J., Kahn, B., Kamm, R. D. & Hwang, W. Atomistic simulation approach to a continuum description of self-assembled beta-sheet filaments. *Biophys. J.* **90**, 2510–2524 (2006).

37. Paparcone, R., Keten, S. & Buehler, M. J. Atomistic simulation of nanomechanical properties of Alzheimer's A beta(1–40) amyloid fibrils under compressive and tensile loading. *J. Biomechanics* **43**, 1196–1201 (2010).

38. Relini, A. *et al.* Detection of populations of amyloid-like protofibrils with different physical properties. *Biophys. J.* **98**, 1277–1284 (2010).

39. Guo, S. & Alshemiteche, B. B. Packing density and structural heterogeneity of insulin amyloid fibrils measured by AFM nanindentation. *Biomacromolecules* **7**, 1630–1636 (2006).

40. Petkova, A. T. *et al.* Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science* **307**, 262–265 (2005).

41. Tanaka, M., Collins, S. R., Toyama, B. H. & Weissman, J. S. The physical basis of how prion conformations determine strain phenotypes. *Nature* **442**, 585–589 (2006).

Role of fibril fragmentation and nanomechanics in the propagation of yeast prion fibrils.

42. Dzwolak, W., Smirnovas, V., Jansen, R. & Winter, R. Insulin forms amyloid in a strain-dependent manner: An FT-IR spectroscopic study. *Protein Sci.* **13**, 1927–1932 (2004).

43. Sigurdson, C. J. *et al.* Prion strain discrimination using luminescent conjugated polymers. *Nature Meth.* **4**, 1023–1030 (2007).

Luminescent conjugated polymers as powerful probes of polymorphism in protein aggregates.

44. Paparcone, R., Cranford, S. W. & Buehler, M. J. Self-folding and aggregation of amyloid fibrils. *Nanoscale* **3**, 1748–1755 (2011).

45. Huang, Y. Y., Knowles, T. P. J. & Terentjev, E. M. Strength of nanotubes, filaments, and nanowires from sonication-induced scission. *Adv. Mater.* **21**, 3945–3948 (2009).

46. Streltsov, V. X-ray absorption and diffraction studies of the metal binding sites in amyloid beta-peptide. *Eur. Biophys. J.* **37**, 257–263 (2008).

47. Ackbarow, T., Chen, X., Keten, S. & Buehler, M. J. Hierarchies, multiple energy barriers and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains. *Proc. Natl. Acad. Sci. USA* **104**, 16410–16415 (2007).

48. Ma, B. & Nussinov, R. Stabilities and conformations of Alzheimer's beta-amyloid peptide oligomers (Abeta 16–22, Abeta 16–35, and Abeta 10–35): Sequence effects. *Proc. Natl. Acad. Sci. USA* **99**, 14126–14131 (2002).

49. Periole, X., Rampioni, A., Vendrusco, M. & Mark, A. E. Factors that affect the degree of twist in beta-sheet structures: a molecular dynamics simulation study of a cross-beta filament of the GNNQQNY peptide. *J. Phys. Chem. B* **113**, 1728–1737 (2009).

50. Lee, C. F., Loken, J., Jean, L. & Vaux, D. J. Elongation dynamics of amyloid fibrils: A rugged energy landscape picture. *Phys. Rev. E* **80**, 041906 (2009).

51. Wei, G. H., Mousseau, N. & Derreumaux, P. Computational simulations of the early steps of protein aggregation. *Proteins* **1**, 3–8 (2007).

52. Xu, Z., Paparcone, R. & Buehler, M. J. Alzheimer's abeta(1–40) amyloid fibrils feature size-dependent mechanical properties. *Biophys. J.* **98**, 2053–2062 (2010).

53. Auer, S. *et al.* Importance of metastable states in the free energy landscapes of poly peptide chains. *Phys. Rev. Lett.* **99**, 178104 (2007).

54. Xu, Z. & Buehler, M. J. Mechanical energy transfer and dissipation in fibrous beta-sheet-rich proteins. *Phys. Rev. E* **81**, 061910 (2010).

55. Fratzl, P. & Weinkamer, R. Nature's hierarchical materials. *Prog. Mater. Sci.* **52**, 1263–1334 (2007).

56. Ashby, M. F., Gibson, I. J., Wegst, U. & Olive, R. The mechanical properties of natural materials. I. Material property charts. *Proc. R. Soc. Lond. A* **450**, 123–140 (1995).

57. Wegst, U. G. K. & Ashby, M. F. The mechanical efficiency of natural materials. *Phil. Mag.* **84**, 2167–2181 (2004).

58. Kieplak, L., Bar, H., Leitner, J. F., Herrmann, H. & Aebi, U. Exploring the mechanical behavior of single intermediate filaments. *J. Mol. Biol.* **354**, 569–577 (2005).

59. Yang, L. *et al.* Micromechanical bending of single collagen fibrils using atomic force microscopy. *J. Biomed. Mater. Res. A* **82**, 160–168 (2007).

60. Shen, Z. L., Dodge, M. R., Kahn, H., Ballarini, S. J., Stress-strain experiments on individual collagen fibrils. *Biophys. J.* **95**, 3956–3963 (2008).

61. Slotta, U. *et al.* Spider silk and amyloid fibrils: A structural comparison. *Macromol. Biosci.* **7**, 183–188 (2007).

62. Vollrath, F. & Knight, D. P. Liquid crystalline spinning of spider silk. *Nature* **410**, 541–548 (2001).

63. Collins, S. R., Donglass, A., Vale, R. D. & Weissman, J. S. Mechanism of prion propagation: amyloid growth occurs by monomer addition. *PLoS Biol.* **2**, e321 (2004).

64. Shorter, J. & Lindquist, S. Destruction or potentiation of different prions catalyzed by similar Hsp104 remodeling activities. *Mol. Cell* **23**, 425–438 (2006).

65. Paparcone, R. & Buehler, M. J. Failure of A-beta-(1–40) amyloid fibrils under tensile loading. *Biomaterials* **32**, 3367–3374 (2011).

Molecular mechanisms of failure of amyloid fibrils and influence of fibril length on mechanical properties.

66. Aguzzi, A. Cell biology: Beyond the prion principle. *Nature* **459**, 924–925 (2009).

67. Prusiner, S. B. Molecular biology of prion diseases. *Science* **252**, 1515–1522 (1991).

68. Riek, R. Cell biology: Infectious Alzheimer's disease? *Nature* **444**, 429–431 (2006).

69. Eisele, Y. S. *et al.* Peripherally applied A beta-containing inoculates induce cerebral beta-amyloidosis. *Science* **330**, 980–982 (2010).

70. Aguzzi, A. & Rajendran, L. The transcellular spread of cytosolic amyloids, prions, and prionoids. *Neuron* **64**, 783–790 (2009).

71. Koffie, R. M. *et al.* Oligomeric amyloid beta associates with postsynaptic densities and correlates with excitatory synapse loss near senile plaques. *Proc. Natl. Acad. Sci. USA* **106**, 4012–4017 (2009).

72. Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nature Rev. Mol. Cell Biol.* **8**, 101–112 (2007).

73. Gebbink, M. F. G., Claessen, D., Bouma, B., Dijkhuizen, L. & Wosten, H. A. B. Amyloids – A functional coat for microorganisms. *Nature Rev. Microbiol.* **3**, 333–341 (2005).

74. Maji, S. K. *et al.* Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325**, 328–332 (2009).

75. Chapman, M. R. *et al.* Role of Escherichia coli curli operons in directing amyloid fiber formation. *Science* **295**, 851–855 (2002).

76. Shorter, J. & Lindquist, S. Prions as adaptive conduits of memory and inheritance. *Nature Rev. Genet.* **6**, 435–450 (2005).

77. Meyer-Lindemann, M. *et al.* Rapid appearance and local toxicity of amyloid-beta plaques in a mouse model of Alzheimer's disease. *Nature* **451**, 720–724 (2008).

78. Barnhart, M. M. & Chapman, M. R. Curli biogenesis and function. *Annu. Rev. Microbiol.* **60**, 131–147 (2006).

79. Hammer, N. D., Schmidt, J. C. & Chapman, M. R. The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. *Proc. Natl Acad. Sci. USA* **104**, 12494–12499 (2007).

80. Zhang, S. G., Holmes, T., Lockshin, C. & Rich, A. Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane. *Proc. Natl Acad. Sci. USA* **90**, 3334–3338 (1993).

81. Zhang, S. G. Fabrication of novel biomaterials through molecular self-assembly. *Nature Biotechnol.* **21**, 1171–1178 (2003).

82. Lovett, M. *et al.* Silk fibroin microtubes for blood vessel engineering. *Biomaterials* **28**, 5271–5279 (2007).

83. Keten, S. & Fuehler, M. J. Geometric confinement governs the rupture strength of H-bond assemblies at a critical length scale. *Nano Lett.* **8**, 743–748 (2008).

84. MacPhee, C. E. & Dobson, C. M. Formation of mixed fibrils demonstrates the generic nature and potential utility of amyloid nanostructures. *J. Am. Chem. Soc.* **122**, 12707–12713 (2000).

85. Reches, M. & Gazit, E. Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **300**, 625–627 (2003).

86. Carmy, O., Shalev, D. E. & Gazit, E. Fabrication of coaxial metal nanotubes using a self-assembled peptide nanotube scaffold. *Nano Lett.* **6**, 1594–1597 (2006).

87. Lu, W. & Lieber, C. M. Nanoelectronics from the bottom up. *Nature Mater.* **6**, 841–850 (2007).

88. Scheibel, T. *et al.* Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proc. Natl Acad. Sci. USA* **100**, 4527–4532 (2003).

89. Niu, L. J., Chen, X. Y., Allen, S. & Tendler, S. J. B. Using the bending beam model to estimate the elasticity of diphenylalanine nanotubes. *Langmuir* **23**, 7443–7446 (2007).

90. Reches, M. & Gazit, E. Controlled patterning of aligned self-assembled peptide nanotubes. *Nature Nanotech.* **1**, 195–200 (2006).

91. Adler-Abanovich, L. *et al.* Self-assembled arrays of peptide nanotubes by vapour deposition. *Nature Nanotech.* **4**, 849–854 (2009).

92. Hamley, I. W. *et al.* Alignment of a model amyloid peptide fragment in bulk and at a solid surface. *J. Phys. Chem. B* **114**, 8244–8254 (2010).

93. Knowles, T. P. J., Oppenheim, T., Buell, A. K., Chirgadze, D. Y. & Welland, M. E. Nanostructured biofilms from hierarchical self-assembly of amyloidogenic proteins. *Nature Nanotech.* **5**, 204–207 (2010).

94. Barrau, S. *et al.* Integration of amyloid nanowires in organic solar cells. *Appl. Phys. Lett.* **93**, 023307 (2008).

95. Channon, K. J., Devlin, G. L. & MacPhee, C. E. Efficient energy transfer within self-assembling peptide fibers: A route to light-harvesting nanomaterials. *J. Am. Chem. Soc.* **131**, 12520–12521 (2009).

96. Liang, Y. *et al.* Light harvesting antenna on an amyloid scaffold. *Chem. Commun.* 6522–6524 (2008).

97. Maji, S. K. *et al.* Amyloid as a depot for the formulation of long-acting drugs. *PLOS Biol.* **6**, e17 (2008).

98. Holmes, T. C. *et al.* Extensive neurite outgrowth and active synapse formation on self-assembling peptide scaffolds. *Proc. Natl Acad. Sci. USA* **97**, 6728–6733 (2000).

99. Ellis-Behnke, R. G. *et al.* Nano neuro knitting: Peptide nanofiber scaffold for brain repair and axon regeneration with functional return of vision. *Proc. Natl Acad. Sci. USA* **103**, 5054–5059 (2006).

**Refs 98 and 99: tissue engineering using amyloid scaffolds.**

100. Gras, S. I. *et al.* Functionalised amyloid fibrils for roles in cell adhesion. *Biomaterials* **29**, 1553–1562 (2008).

101. Gimona, M. Protein linguistics – a grammar for modular protein assembly? *Nature Rev. Mol. Cell Biol.* **7**, 68–73 (2006).

102. Diaz-Avalos, R. *et al.* Cross-beta order and diversity in nanocrystals of an amyloid-forming peptide. *J. Mol. Biol.* **330**, 1165–1175 (2003).

103. Hemingway E. *The Old Man and the Sea* (Vintage Books, 2007).

104. Nelson, R. *et al.* Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).

105. Galkin, V. E., Orlova, A., Cherepanova, O., Lebart, M.-C. & Ebelman, E. H. High-resolution cryo-EM structure of the F-actin-fibrin/plastin ABD2 complex. *Proc. Natl Acad. Sci. USA* **105**, 1494–1498 (2008).

106. Li, H., DeRosier, D. J., Nicholson, W. V., Nogales, E. & Downing, K. H. Microtubule structure at 8 Å resolution. *Structure* **10**, 1317–1328 (2002).

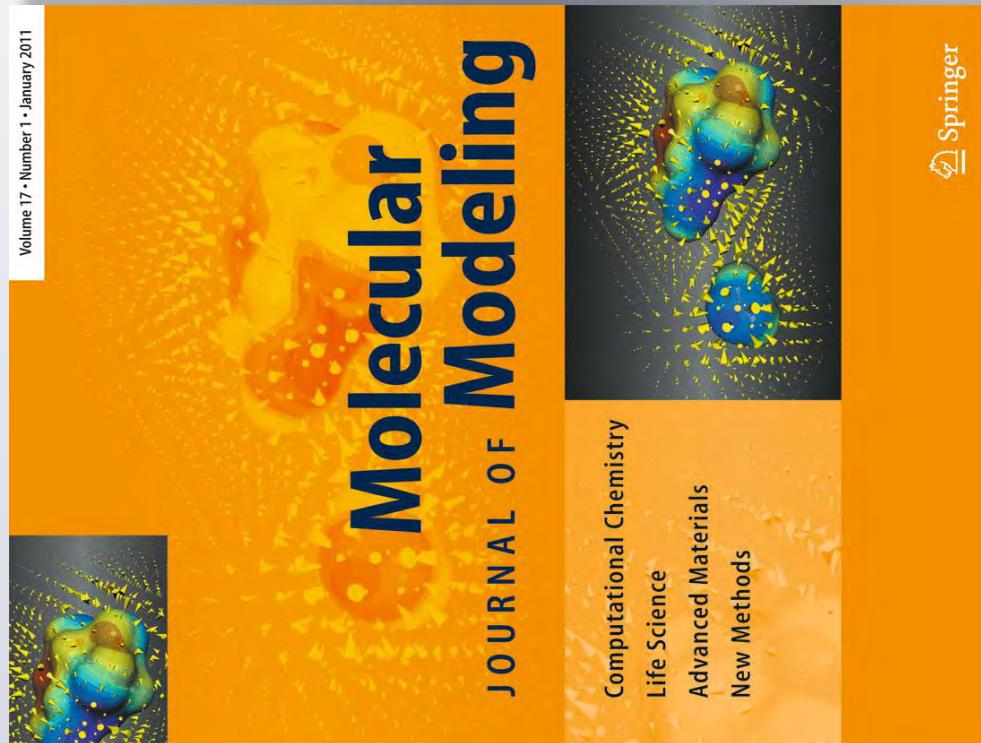
## Acknowledgements

T.P.J.K. acknowledges support from St John's College, Cambridge. M.J.B. acknowledges support from the Office of Naval Research (YIP and PECASE Awards), National Science Foundation (CAREER), Army Research Office and the Air Force Office of Scientific Research. We also acknowledge helpful discussions with D. Kaplan, A. Aguzzi, L. Luheshi, D. White and C. Dobson.

## Additional information

The authors declare no competing financial interests.

*Structure and dynamics of human vimentin intermediate filament dimer and tetramer in explicit and implicit solvent models*



**Journal of Molecular**

**Modeling**

Computational Chemistry - Life  
Science - Advanced Materials -  
New Methods

ISSN 1610-2940

Volume 17

Number 1

J Mol Model (2010) 17:37–48  
DOI 10.1007/  
s00894-010-0696-6

Computational Chemistry  
Life Science  
Advanced Materials  
New Methods

 Springer

 Springer

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

# Structure and dynamics of human vimentin intermediate filament dimer and tetramer in explicit and implicit solvent models

Zhao Qin · Markus J. Buehler

Received: 22 December 2009 / Accepted: 22 February 2010 / Published online: 1 April 2010  
© Springer-Verlag 2010

**Abstract** Intermediate filaments, in addition to microtubules and microfilaments, are one of the three major components of the cytoskeleton in eukaryotic cells, and play an important role in mechanotransduction as well as in providing mechanical stability to cells at large stretch. The molecular structures, mechanical and dynamical properties of the intermediate filament basic building blocks, the dimer and the tetramer, however, have remained elusive due to persistent experimental challenges owing to the large size and fibrillar geometry of this protein. We have recently reported an atomistic-level model of the human vimentin dimer and tetramer, obtained through a bottom-up approach based on structural optimization *via* molecular simulation based on an implicit solvent model (Qin et al. in PLoS ONE 2009, 4(10):e7294, 9). Here we present extensive simulations and structural analyses of the model based on ultra large-scale atomistic-level simulations in an explicit solvent model, with system sizes exceeding 500,000 atoms and

simulations carried out at 20 ns time-scales. We report a detailed comparison of the structural and dynamical behavior of this large biomolecular model with implicit and explicit solvent models. Our simulations confirm the stability of the molecular model and provide insight into the dynamical properties of the dimer and tetramer. Specifically, our simulations reveal a heterogeneous distribution of the bending stiffness along the molecular axis with the formation of rather soft and highly flexible hinge-like regions defined by non-*alpha*-helical linker domains. We report a comparison of Ramachandran maps and the solvent accessible surface area between implicit and explicit solvent models, and compute the persistence length of the dimer and tetramer structure of vimentin intermediate filaments for various subdomains of the protein. Our simulations provide detailed insight into the dynamical properties of the vimentin dimer and tetramer intermediate filament building blocks, which may guide the development of novel coarse-grained models of intermediate filaments, and could also help in understanding assembly mechanisms.

---

Z. Qin · M. J. Buehler (✉)

Laboratory for Atomistic and Molecular Mechanics,  
Department of Civil and Environmental Engineering,  
Massachusetts Institute of Technology,  
77 Massachusetts Ave., Room 1-235A&B,  
Cambridge, MA, USA  
e-mail: mbuehler@MIT.EDU

Z. Qin · M. J. Buehler  
Center for Materials Science and Engineering,  
Massachusetts Institute of Technology,  
77 Massachusetts Ave.,  
Cambridge, MA, USA

M. J. Buehler

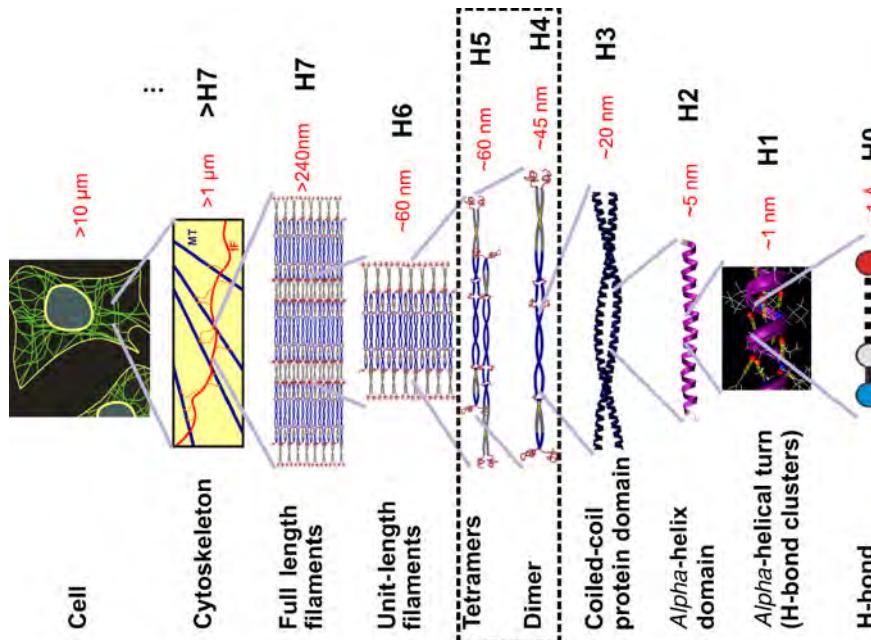
Center for Computational Engineering,  
Massachusetts Institute of Technology,  
77 Massachusetts Ave.,  
Cambridge, MA, USA

## Introduction

Intermediate filaments (IFs), in addition to microtubules and microfilaments, are one of the three major components of the cytoskeleton in eukaryotic cells [1] and are emerging as major players in cell biology, in particular in issues related to mechanobiological mechanisms. IFs are crucial in

defining key mechanical functions of cells such as cell migration, cell division and mechanotransduction, and have also been referred to as the “safety belts of cells” in the biology community as they play a structural role in preventing exceedingly large cell stretch [2–6]. Earlier studies focused on analyzing the mechanical signature of IFs have suggested that they are highly sensitive to applied forces at small levels, and that they can sustain extremely large deformation of up to 300% or four times their initial length [7, 8], an observation that was also observed in recent simulation studies [9]. Proteins of the IF family are also found in the cell’s nucleus in the form of lamin, where they form a dense mesh-like network providing mechanical integrity and biochemical functions at the cytoskeleton–chromatin interface [10–12]. IFs have been associated with many genetic diseases, where single point mutations, domain deletions as well as sequence modifications lead to structural changes at different levels in the IFs’ hierarchical organization. IF related diseases include muscle dystrophies, Alexander disease [13], epidermolysis bullosa simplex, as well as a broad class of disorders referred to as laminopathies (e.g., rapid aging disease progeria) [14–17]. In extracellular materials, IF proteins are found in hair, hoof and wool, further emphasizing the broad significance of this protein material for a variety of applications in biology, medicine and engineering.

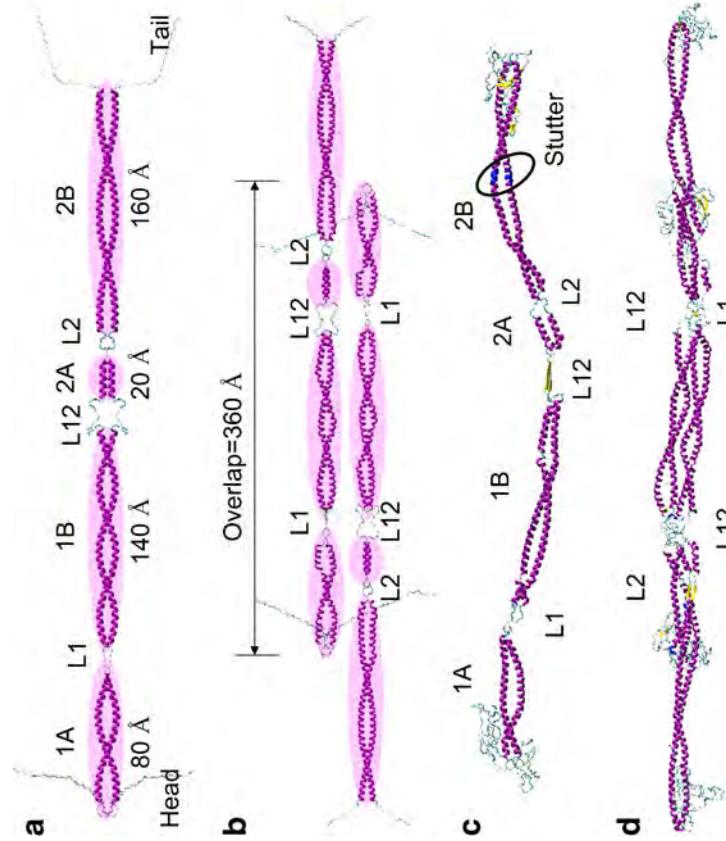
IFs form hierarchical structures as shown in Fig. 1, ranging from dimers, tetramers, unit length filaments, full-length filaments to the cellular level. Each vimentin dimer, the basic building block of this protein filament, contains 466 amino acids. Experimental studies suggested that they form four major structural segments linked in series in the sequence 1A, 1B, 2A and 2B, connected by linkers L1, L12 and L2 (see Fig. 2a). Up until now, a complete atomistic-level molecular model of the dimer or tetramer structure has remained elusive. So far, only parts of two of the four segments of the dimer, a section of the 1A and 2B domain of the vimentin dimer structure, have been crystallized and their atomic structure identified based on x-ray diffraction experiments [18–22] (the structures are found in Protein Data Bank (PDB) entries 1gk4, 1gk6, 3klt and 1gk7). Persistent experimental challenges have prevented the identification of the remaining parts of the IF structure using experimental approaches. This is because IFs are intrinsically disordered structures [23], so that x-ray diffraction studies on naturally occurring or recombinantly produced IF bundles do not provide sufficient data to produce a full atomistic model of IFs. Whereas solid state nuclear magnetic resonance (ssNMR) has been successfully utilized to derive atomistic models of amyloid fibrils [24], this approach has not yet worked for IF dimers or tetramer, likely because amyloid peptides are much smaller than IF dimers. Cryoelectron tomography has been suggested as a



**Fig. 1** The hierarchical structure of intermediate filaments, from atomic to cellular scales. The figure shows relevant structural levels (H0 to H7). The present paper focuses on structural analysis of levels H4 and H5 (dimer/tetramer structure). The regions marked in yellow in the dimer level are those that have been crystallized in earlier studies. The focus of the study reported in this paper is on vimentin dimers and tetramers, corresponding to hierarchy levels H4 and H5 as marked in the figure. Figure adapted from reference [9]

possible molecular-level imaging tool to visualize single IFs, but the highest resolution tomograms are still limited to a resolution around 5 nm [25].

The lack of a complete structural model of IFs has prevented us from addressing fundamental structure–function relationship questions related to the mechanical role of intermediate filaments, which is crucial to link structure and function in the protein filament’s biological context. The availability of a structural model of IFs could also be the key to understanding the mechanisms of IF related genetic diseases, where structural flaws that originate at the genetic or molecular level may cause major structural changes of biologically relevant properties. To address these issues, here we provide a detailed analysis of a molecular model of IF dimers and tetramers with atomistic resolution that was recently obtained based on a bottom-up molecular simulation approach using an effective solvent model [9]. The earlier structure identification



**Fig. 2** Schematics of the molecular structure of the vimentin dimer (panel **a**) and the vimentin tetramer (panel **b**). The plot includes labels identifying the various segments and linker domains, as well as the head and tail domain. Panels **c** and **d** show the equilibrated structures of the dimer and tetramer, obtained from implicit solvent simulations as reported in [9]

approach utilized the available amino acid sequence and associated structural information from experimental results to generate an initial geometry that served as the starting point for a series of energy minimization and equilibration runs, which were subsequently performed using molecular dynamics. These simulations resulted in an equilibrated structure of both the vimentin dimer and tetramer.

Since the earlier study focused primarily on simulations in implicit solvent, an important outstanding question remains a detailed structural and dynamical study of the behavior of the resulting proteins in explicit solvent. Several earlier comparisons between explicit and implicit solvents have been reported in the literature, focused on exploring the effects of the solvents in conformation sampling and energy landscape in folding [26, 27]. In these studies it was found that simulation of proteins in implicit solvent can reproduce the set of local energy minimum in explicit solvent, but some energy terms such as salt bridge may be overestimated. However, most of the earlier works were focused on small peptides. The study reported here compares the effects of implicit versus explicit solvent for the cases of a IF dimer and tetramer, which are much larger proteins than the small peptides studied before (932 amino acids of a dimer and 1864 amino acids of a tetramer). To address this issue, here we present large-scale validation simulations and structural analyses of the atomistic model using structural equilibration in large-scale atomistic simulations in explicit solvent, with system sizes up to 500,000 atoms and simulations carried out at 20 ns time-scales. Based on these explicit solvent simulations, we report a

detailed comparison of the dynamical behavior of this biomolecular model within implicit and explicit solvent models. Our studies confirm the stability of the molecular model, and provide insight into the dynamical properties of the dimer and tetramer.

#### Materials and methods

Background information on implicit solvent versus explicit solvent models

The solvent environment plays an important role in the structure, dynamics, and function of biomolecules. It is generally agreed upon that an explicit solvent model provides the most accurate treatment of solute-solvent interaction [28], because each explicit solvent molecule reflects its realistic chemical structure and resulting interactions with the solute and one another. However, the approach of relying on an explicit treatment of solvent molecules significantly increases the system size by several orders of magnitude. To achieve a better balance of computational efficiency and physical accuracy, implicit solvent models have been proposed as an alternative approach [29, 30]. Implicit solvent models add the effective solvent energy to the biological molecule, expressed indirectly as a function of the molecular structure and its conformation.

This strategy simplifies the numerical effort of simulating the physical system of interest and thereby speeds up

the simulation efficiency by averaging the solvent effect. Since no frictional effects are included in implicit solvent methods, such models can facilitate a more rapid structural organization toward the equilibrium geometry due to the lack of viscosity imposed by explicit water molecules. This effect can be advantageous in structure identification approaches [31–33] (such as in the study reported in [9]) since it accelerates the sampling of molecular configurations, so that more configurations are visited per simulated unit time. Whereas the approach of using implicit solvents is computationally efficient, it has been discussed that the implicit model may overestimate some energy terms and lead to inaccurate results, in particular regarding the detailed configurations of domains in a protein [27]. Therefore, further validation and comparison by explicit solvent models is needed to ensure the validity of the results obtained from implicit solvent simulations. To address these issues, a detailed comparison between implicit and explicit simulation models is provided here, applied specifically to a vimentin intermediate filament dimer and tetramer.

#### Molecular models and force field

The implicit solvent simulations whose results are analyzed here [9] were carried out with the CHARMM program by using the CHARMM19 all-atom energy function and an implicit Gaussian model for the water solvent [29, 30]. Explicit solvent simulations [34] are used to confirm the stability of the equilibrated structures, showing that the structures predicted by our approach are stable also in explicit solvent. We use the CHARMM force field and explicit TIP3 water [34] implemented in NAMD [35], starting from the equilibrated structure obtained using the approach described above. The time step used in all simulations is 1 fs, at a constant temperature (300 K) controlled by Langevin thermostat. A pressure of 1 atmosphere is used in the explicit water simulations. We use the particle mesh Ewald (PME) method in calculating the electrostatic interactions, with a grid size of 1 Å in all three directions. A cutoff length of 10 Å is applied to the van der Waals interactions and to switch the electrostatic forces from short range to long range forces. We use pure water in our simulations (no ions are included).

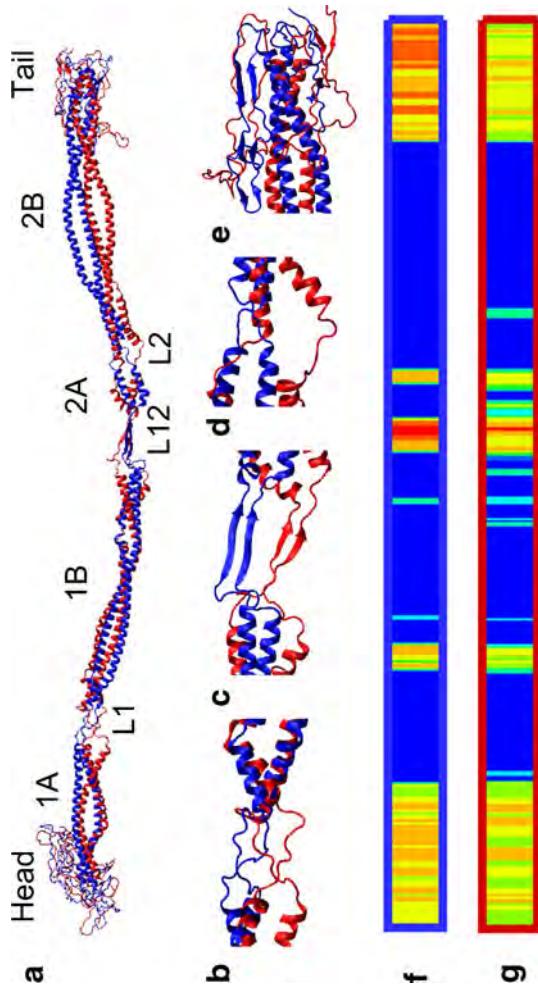
#### Initial molecular structure

The structure of the dimer has been identified by a series of computational steps, as reported in [9]. Figure 2a and b show a schematic of the structural character of the dimer and the tetramer, with each domain labeled. Figure 2c and d depict a snapshot of the equilibrated dimer and tetramer structures as obtained from implicit solvent calculations.

These geometries from earlier studies [9] are used as a starting point for the simulations reported here. The dimer structure is reported at <http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info%3Adoi%2f10.1371%2fjournal.pone.0007294.s010>, and the tetramer structure is reported at <http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info%3Adoi%2f10.1371%2fjournal.pone.0007294.s011> is used. Explicit water is added in a rectangular water box with periodic boundary conditions in all three directions using visual molecular dynamics (VMD). We use a rectangular solvent box with dimension of 54 nm × 7.5 nm × 7 nm for the dimer and 66 nm × 9 nm × 9 nm for the tetramer. Because the diameters of the dimer and tetramer are estimated to be 2.4 nm and 4 nm, respectively, the nearest distances between two neighboring dimers and tetramers are 4.6 nm and 5 nm, respectively. These distances are much larger than the cutoff length of the pair list distance (12 Å) and thereby there is no interaction between the neighboring dimers or tetramers. The dimer and tetramer systems in explicit solvent contain ≈250,000 in the dimer and ≈500,000 atoms in the tetramer case (in this case ≈30,000 atoms for the protein plus ≈470,000 for the water molecules). The use of parallelized supercomputers enables us to reach about 20 ns in about one month of simulation time.

#### Results

We begin with a comparison of the equilibrated vimentin dimer and tetramer geometries as obtained from implicit and explicit equilibration. Figure 3a shows snapshots of the molecular structure of the vimentin dimer, displaying the characteristic segmented geometry with coiled-coil regions connected through linker domains. The figure presents a comparison of the structure based on implicit solvent equilibration and explicit solvent equilibration, where the two geometries are superimposed in the plot. Figure 3b–e show detailed views of key domains in the protein structure, also in a visualization mode where the implicit and explicit solvent model results are superimposed. It can be seen that the linker L12 forms a parallel *beta*-sheet in both explicit and implicit solvent models. We find that in both cases the total length of the dimer without the head and tail domain is ≈49 nm. This result is in agreement with experimental results where a range of 46–49 nm has been reported [19, 36]. The visual comparison of the dimer structure between implicit and explicit solvent models shows that the overall geometry of the dimer is stable in either solvent model. We present a detailed analysis of the structural character of the equilibrated structure of the dimer in the two solvents in Fig. 3f and g, where the color along the molecular axis

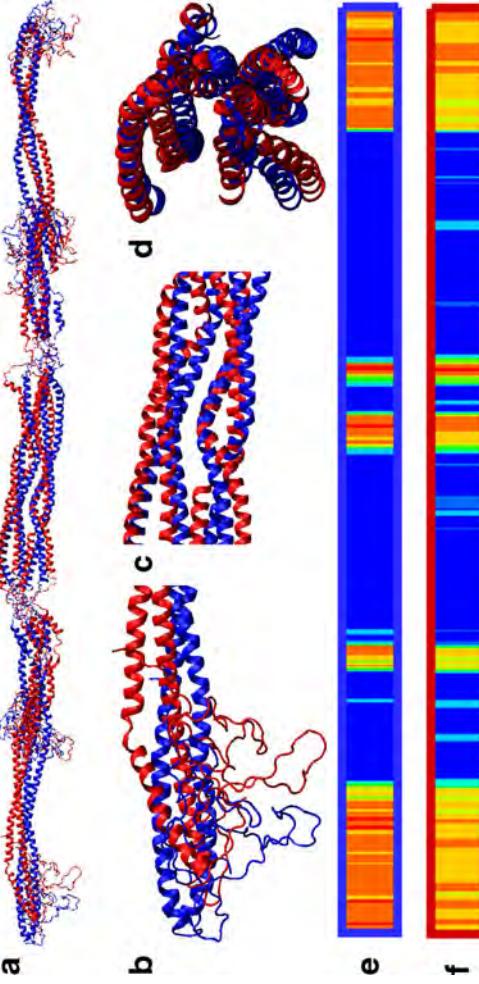


**Fig. 3** Structure of the vimentin intermediate filament dimer, depicting a comparison between implicit and explicit solvent simulation results. Panel **a**: Snapshot of the atomistic model of the dimer after equilibration, in implicit water (in blue) and in explicit water (in red). Panels **b–e** show detailed views of the structure of linkers L1, L12, L2 and the tail domain, respectively. The *beta*-sheet structure of linker L12 is visible in panel **c** (we note that while the

denotes the protein filament's secondary structure. By comparing the two structures, we find that the 2A segment (parallel helices) and L12 (*beta*-sheet) are only marginally stable, because the 2A segment lacks the hierarchical structure in enforcing its strength, and short *beta*-sheet segments are more energy favorable at the more flexible parts of the protein [37].

*beta*-sheet is rather stable in the implicit simulations, it is only marginally stable in the explicit solvent model). Panels **f** and **g** show the structural character of each amino acid along the polypeptide chain length, for the implicit (panel **f**) and explicit (panel **g**) solvent. *Alpha*-helical structures are colored in blue, *beta*-sheet structures are colored in red, and the other colors represent random coils

Figure 4a shows a comparison of the vimentin tetramer structure of the implicit solvent equilibration and explicit solvent equilibration results, where the two geometries are superimposed in the plot. The total length of the tetramer is  $\approx 60$  nm with a slightly smaller length of  $\approx 59$  nm in the explicit solvent case. These findings agree well with the length of unit length filaments observed in experiment,



**Fig. 4** Structure of the vimentin intermediate filament tetramer, depicting a comparison between implicit and explicit solvent simulation results. Panel **a**: Snapshot of the atomistic model of the dimer after equilibration, in implicit water (in blue) and in explicit water (in red). Panels **b, c, d** and **e** show detailed views of the structure of head and overlapped domain from two viewing

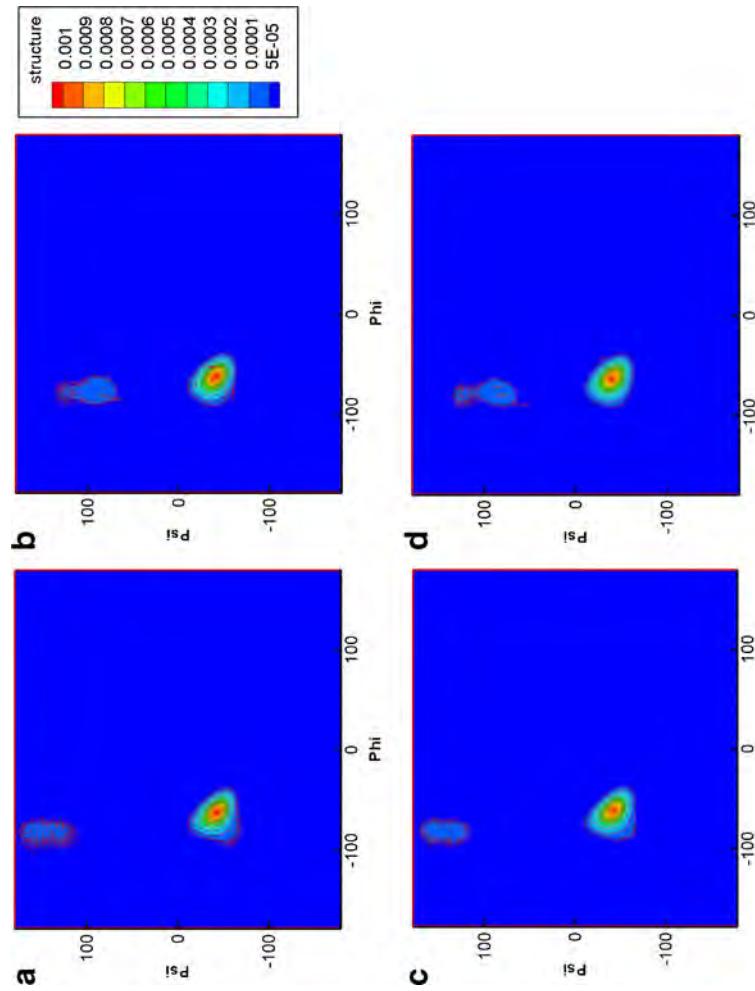
directions, respectively. Panels **e** and **f** show the structural character of each amino acid along the polypeptide chain length, for the implicit (panel **e**) and explicit (panel **f**) solvent. *Alpha*-helical structures are colored in blue, *beta*-sheet structures are colored in red, and the other colors represent random coils

62 nm at pH 7.5 [36]. In both cases the overlap part of the tetramer has a length of  $\approx 36$  nm, where the experimental value is 30–36 nm [19, 36]. Segments 1A, 1B and 2A are fully contained in this overlapped part, but the 2B segments in the immediate vicinity of the two terminals are located completely outside. We find that the head segment of each dimer is coiled around the other dimer, increasing the contact surface area and thereby providing enhanced interdimer interactions, as shown in Fig. 4b. The structure of the coiled-coil dominated overlap region is rather stable, as shown in Fig. 4c and d (the images show the same part of the tetramer from different viewing directions). As for the dimer case discussed in the previous paragraph, here we also present a detailed analysis of the structural character of the equilibrated protein in Fig. 4e and f. The analysis shows that in comparison with the dimer case, the tetramer structure is more stable, and most of the structural features are well conserved after explicit solvent equilibration.

We analyze the structure of the dimer and tetramer by calculating the dihedral angle distribution as shown

in the Ramachandran map [38] as depicted in Fig. 5. This analysis shows the probability of specific dihedral angles  $[\varphi, \psi]$  of all amino acid residues in the dimer/tetramer structure in equilibrium. The analysis shows that most of the residues ( $\approx 73\%$ ) belong to the left lower corner ( $\varphi < 0^\circ, \psi < 0^\circ$ ), and the peak value emerges at  $[\varphi \approx -58^\circ, \psi \approx -47^\circ]$ . This data shows the character of a typical right-handed *alpha*-helix. Most other residues ( $\approx 25\%$ ) belong to the left upper corner ( $\varphi < 0^\circ, \psi > 0^\circ$ ), and there is a much less intense peak that resembles the character of a *beta*-sheet in a Ramachandran map ( $-180^\circ < \varphi < -50^\circ$  and  $50^\circ < \psi < 180^\circ$ ). The position of this peak center for the dimer/tetramer structure in implicit solvent ( $[\varphi \approx -80^\circ, \psi \approx 100^\circ]$ ) is slightly different from that of the explicit solvent ( $[\varphi \approx -80^\circ, \psi \approx 140^\circ]$ ), indicating that the *beta*-sheet in explicit solvent has a more extended conformation than that in implicit solvent.

We further calculate the number of H-bonds in the dimer and tetramer structure during equilibration, and compare the data based on implicit and explicit solvent simulations as



**Fig. 5** Ramachandran maps that show the averaged dihedral angle of each amino-acid  $([\varphi, \psi])$  distribution during the equilibrium process, comparing implicit and explicit simulation results. Panels **a** and **b** show the results for the dimer in explicit solvent (**a**) and implicit solvent (**b**). Panels **c** and **d** show the results for the tetramer in explicit solvent (**c**) and implicit solvent (**d**). The analysis shows that most of the residues ( $\approx 73\%$ ) belong to the left lower corner ( $\varphi < 0^\circ, \psi < 0^\circ$ ), and the peak value emerges at  $[\varphi \approx -58^\circ, \psi \approx -47^\circ]$ . This data shows the characteristic of a right-handed *alpha*-helix. Most

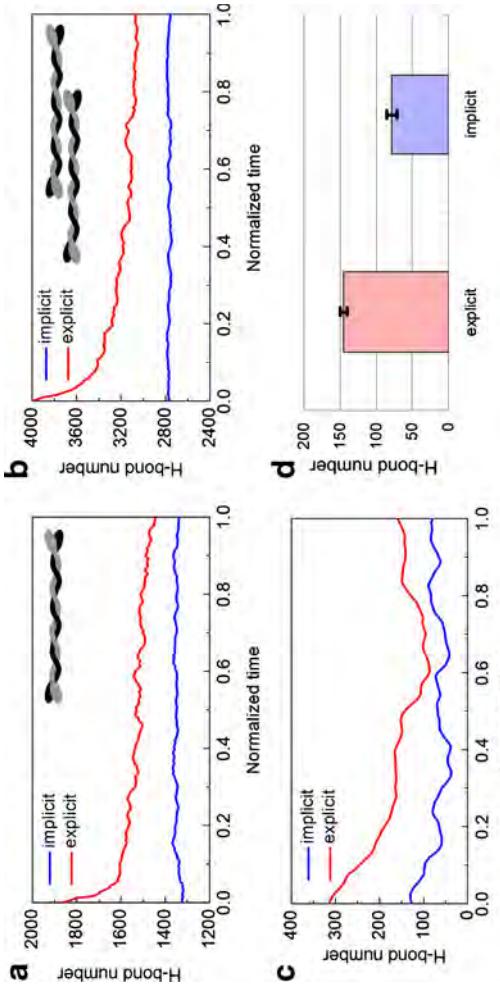
other residues ( $\approx 25\%$ ) belong to the left upper corner ( $\varphi < 0^\circ, \psi > 0^\circ$ ), and there is a much less intense peak that resembles the character of a *beta*-sheet in a Ramachandran map ( $-180^\circ < \varphi < -50^\circ$  and  $50^\circ < \psi < 180^\circ$ ). The position of this peak center for the dimer/tetramer structure in implicit solvent ( $[\varphi \approx -80^\circ, \psi \approx 100^\circ]$ ) is slightly different from that of the explicit solvent ( $[\varphi \approx -80^\circ, \psi \approx 140^\circ]$ ), indicating that the *beta*-sheet in explicit solvent has a more extended conformation than that in implicit solvent

shown in Fig. 6. We observe that the number of H-bonds in both the dimer and tetramer structure decreases after energy minimization and converges asymptotically to a constant value in explicit solvent. The asymptotic number of H-bonds is also stable for the simulations in implicit solvent, as shown in Fig. 6a and b. By subtracting two times the number of H-bond in the dimer from that of the tetramer, we estimate the number of *inter*-dimer H-bonds, which varies with time as shown in Fig. 6c. The asymptotic value by the end of the simulation (by averaging over the last 20% of the total simulation time) is given in Fig. 6d, where it is observed that the number of *inter*-dimer H-bonds in explicit solvent ( $145 \pm 5$ ) is greater than that obtained in implicit solvent ( $78 \pm 8$ ), indicating that the interaction between two dimers in explicit solvent might be stronger than that of the implicit solvent.

The solvent accessible surface area (SASA), which reflects the surface area that is accessible to a water molecule with radius of 1.4 Å, of the dimer and tetramer is shown in Fig. 7a and b. It is shown that the SASA in the dimer and tetramer structures in explicit solvent increases after energy minimization, while it is relatively stable in implicit solvent. This phenomenon agrees with the observations made in the  $[\phi, \psi]$  analysis, and suggests that the molecular conformation in explicit solvent tends to become more extended leading to an increase in the SASA value. We calculate the covered *inter*-dimer surface area by subtracting the SASA value of the tetramer from two times of the SASA value of the dimer, as shown in Fig. 7c. The

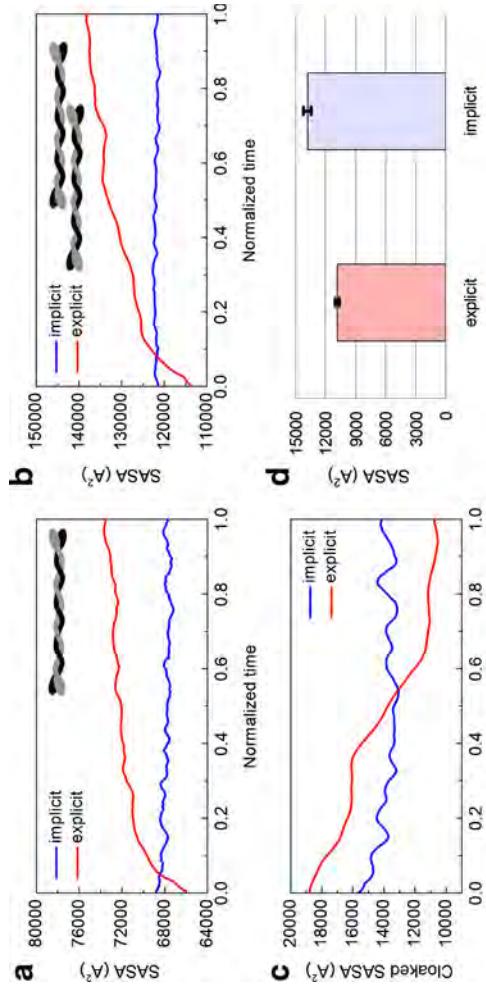
asymptotic value by the end of the simulation (by averaging over the last 20% of the total simulation time) is shown in Fig. 7d, where it can be seen that the two dimers in a tetramer model become less adhesive in explicit solvent as they allow for more water to fill the gap between the dimers as shown in Fig. 7c. This phenomenon is not observed in implicit solvent. By the end of the simulation, the hidden surface areas of the implicit and explicit solvent simulations converge to  $13,800 \text{ \AA}^2$  and  $10,800 \text{ \AA}^2$ , respectively, as shown in Fig. 7d. This finding suggests that the hidden area between two dimers in explicit solvent is 22% smaller than that observed under implicit solvent conditions.

From dynamical analyses of the trajectory of the dimer at 300 K, we observe that the linker domains represent rather soft, hinge-like structures that connect much stiffer coiled-coil segments. The heterogeneous distribution of the bending stiffness along the molecule's axis strongly affects the filaments' motion by enhancing its mobility due to an effectively reduced persistence length, a mechanism that might be significant for IF self-assembly mechanisms. Thereby, the linker domains act as hinges around which the molecular structure can more easily bend and rotate. The particular distribution of soft and stiff elements along the dimer axis suggests a simplified model of the protein structure as shown in Fig. 8a. In this representation, each coiled-coil is represented by a rod and each linker is represented by a hinge. Thereby the distances  $L1$ ,  $L2$ ,  $L3$  and  $L4$ , as well as the angles  $\theta1$ ,  $\theta2$  and  $\theta3$  describe the key geometrical parameters of the dimer structure. We



**Fig. 6** Number of H-bonds of the vimentin intermediate filament dimer and tetramer, showing a comparison between implicit and explicit solvent simulation results. Panel **a**: Number of H-bonds in a dimer. Panel **b**: Number of H-bonds in a tetramer. Panel **c**: Number of H-bonds between the two dimers within a tetramer. Each of the panels **a**, **b** and **c** are obtained during the equilibration process, in both implicit and explicit solvent, respectively. Panel **d**: Comparison between the asymptotic values of the number of interdimer H-bonds

in both implicit and explicit solvent. The main bar gives the average value during the last 20% of the total simulation time while the corresponding error bar corresponds to the standard deviation. In panels **a**–**c**, the simulation time is normalized by the length of the total equilibration (10 ns for the implicit solvent for both dimer and tetramer, 16 ns for the explicit solvent for the dimer, and 19 ns for explicit solvent simulations for the tetramer)



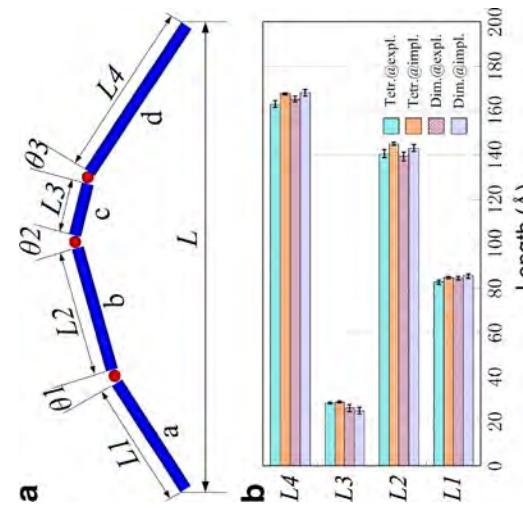
**Fig. 7** Solvent accessible surface area (SASA) of the vimentin intermediate filament dimer and tetramer, showing a comparison between implicit and explicit solvent models. Panel **a**: SASA of the dimer. Panel **b**: SASA of the tetramer. Panel **c**: Cloaked SASA between the two dimers within a tetramer. Each of the Panel **a**, **b** and **c** report results obtained during the equilibrium process, in implicit and explicit solvent, respectively. Panel **d**: Comparison between the asymptotic values of the between-dimer-SASA in implicit and

measure the value of the distances  $L1$ ,  $L2$ ,  $L3$  and  $L4$ , as well as the angles  $\theta1$ ,  $\theta2$  and  $\theta3$  during both the implicit and explicit equilibration and summarize the resulting mean and standard deviation values in Table 1. Figure 8b

compares the length of each rod-like domain in both the dimer and the tetramer based on the implicit and explicit solvent model. It is shown that the lengths of all the four coiled-coil domains remain almost constant during simulation, providing evidence that both implicit and explicit models provide a similar description of the structure of the rod-like domains. The geometric location of the bending angles  $\theta1$ ,  $\theta2$  and  $\theta3$  as a function of the normalized simulation time are shown in Fig. 9a–c. We observe that the bending angle of each linker is fluctuating around an average value during the simulation. This average value and the fluctuation range is summarized in Fig. 9d, in which we plot the mean value and standard deviation of the three angles. We observe that the fluctuations of the angles in the tetramer are smaller than in the dimer, indicating that the structure is stabilized in the assembling process. Using the definition of the persistence length  $L_p$

$$\cos(\theta_{mean}) = \left\langle \vec{t}(s) \cdot \vec{t}(s') \right\rangle = \exp \left[ -\frac{(s - s')}{L_p} \right] \quad (1)$$

where  $s$  and  $s'$  are coordinates along the contour lengths of the chain. Using the mean angle values in the tetramer obtained from the explicit solvent simulations we compute the persistence length of the three parts, each of which has a linker at its center, as  $L_p^{ab} = 322\text{nm}$ ,  $L_p^{bc} = 234\text{nm}$  and  $L_p^{cd} = 392\text{nm}$ . The effective persistence length of a full dimer/tetramer in different solvents (implicit/explicit). The mean value is given by the main bar, and standard deviations are plotted *via* error bars



**Fig. 8** Analysis of the geometric parameters of a dimer. Panel **a**: Simplified dynamical model of a dimer, which is considered here as a collection of rod-like domains connected by hinge-like linkers. The points a, b, c and d correspond to the mid points of the corresponding rod-like domains. Panel **b**: the length of each rod-like domain within dimer/tetramer in different solvents (implicit/explicit). The mean value is given by the main bar, and standard deviations are plotted *via* error bars

**Table 1** Structure analysis of each domain within the dimer and tetramer in implicit solvent and explicit solvent during equilibrium, the geometric parameters are as shown in Fig. 5a

Geometric parameter	Dimer in implicit solvent	Dimer in explicit solvent	Tetramer in implicit solvent	Tetramer in explicit solvent
$L1$	Average (Å)	85.5	84.6	84.8
	SD. $\pm$ (Å)	1.0	0.9	0.5
$L2$	Average (Å)	143.2	139.6	145.1
	SD. $\pm$ (Å)	1.6	2.0	0.8
$L3$	Average (Å)	24.8	26.1	28.8
	SD. $\pm$ (Å)	1.5	1.5	0.5
$L4$	Average (Å)	168.0	165.2	167.5
	SD. $\pm$ (Å)	1.5	1.4	0.5
$L$	Average (Å)	487.0	479.7	613.0
	SD. $\pm$ (Å)	—	2.2	—
$\theta1$	Average (°)	10.7	12.0	15.8
	SD. $\pm$ (°)	4.7	5.7	2.9
$\theta2$	Average (°)	13.9	12.7	7.9
	SD. $\pm$ (°)	6.9	6.0	4.4
$\theta3$	Average (°)	13.7	11.6	11.2
	SD. $\pm$ (°)	7.1	6.3	4.1

It is noted that except for  $L$  (which represents the end-to-end length of a dimer or tetramer), all other parameters are measured based on the rod-like domains and linkers as identified in the dimer structure. SD = standard deviation, used here to define error bars

Based on the estimated persistence length, the bending modulus of the dimer can be expressed as [39]

$$E_{bend} = k_B T L_p / I \quad (2)$$

where  $I = \frac{\pi d^4}{64}$ , and  $d = 2.4$  nm is the diameter of the dimer including the vdW radius (obtained from geometric analysis of the protein structure). Thereby, the bending modulus is found to be  $E_{bend} = 1$  GPa. The bending modulus obtained from experimental bending was reported to be around 900 MPa in earlier studies [39], close to our results.

In extending this analysis, we estimate the angular stiffness of each hinge shown in Fig. 8a. Using a beam theory for a clamped beam, the angle change is given by:

$$\Delta\theta = \frac{Ml}{E_{bend}I} \quad (3)$$

where  $l$  is the contour length of the domain considered (for example, in studying  $\overline{d\theta}$ , we have  $l = (l_1 + l_2)/2$ ). Since  $K_\theta = \frac{M}{\Delta\theta}$ , combining with Eq. 2 and we obtain

$$K_\theta = \frac{k_B T L_p}{l} \quad (4)$$

Based on this analysis the angular stiffness for each of the three hinges is calculated to be  $K_{\theta1} = 1.72$  kcal/mol (for L1),  $K_{\theta2} = 1.65$  kcal/mol (for L12), and  $K_{\theta3} = 2.44$  kcal/mol (for L2). It is observed that all these values are relatively small, which explains why each rod-like

domain can easily rotate around its linker, in agreement with the observations made in the dynamical simulations.

Finally, Fig. 10 shows the evolution of the root mean square deviation (RMSD) for the dimer (Fig. 10a) and the tetramer (Fig. 10b) in explicit solvent (compared with the initial structure as obtained from implicit solvent equilibration). The results show that after 12 ns, the RMSD values for both the dimer and tetramer converge to a constant value, suggesting that both structures have reached a stable configuration in the explicit solvent model (the final structures are the structures shown in Figs. 3 and 4, respectively). Furthermore, the overall relatively small deviation of the structure provides evidence for the fact that the implicit model is a reasonable approximation for the dimer and tetramer structure, at least at the time-scales considered here.

Additional analysis of the dynamical properties of each amino-acid in the dimer and tetramer during the simulation is shown in Fig. 11, where the RMSD value for each residue in the simulation is depicted. These results also reflect the relative flexibility of each domain during the equilibration process. We observe that the peaks shown in Fig. 11a correspond to the linker, head and tail segments, providing the evidence that these domains are among the most flexible ones in the dimer structure and as such undergo the largest structural fluctuations. The peak RMSD values  $P$  corresponding to the linkers have a relation as  $P_{L12} > P_{L1} > P_{L2}$ , which agrees with the relation of the angular stiffness of each linker  $K_{\theta2}K_{\theta1}K_{\theta3}$ . It is also



## Structure and stability of the lamin A tail domain and HGPS mutant

Zhao Qin <sup>a,1</sup>, Agnieszka Kalinowski <sup>b,1</sup>, Kris Noel Dahl <sup>b,c,\*</sup>, Markus J. Buehler <sup>a,\*</sup>

<sup>a</sup> Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77

Mass. Ave., Room 1-235a&B, Cambridge, MA 02139, USA

<sup>b</sup> Department of Biomedical Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

<sup>c</sup> Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

### ARTICLE INFO

#### Article history:

Received 8 March 2011

Received in revised form 13 May 2011

Accepted 17 May 2011

Available online 24 May 2011

### ARTICLE

Hutchinson–Gilford progeria syndrome (HGPS) is a premature aging syndrome caused by the expression and accumulation of a mutant form of lamin A,  $\Delta 50$  lamin A. As a component of the cell's nucleoskeleton, lamin A plays an important role in the mechanical stabilization of the nuclear envelope and in other nuclear functions. It is largely unknown how the characteristic 50 amino acid deletion affects the conformation of the mostly intrinsically disordered tail domain of lamin A. Here we perform replica exchange molecular dynamics simulations of the tail domain and determine an ensemble of semi-stable structures.

Based on these structures we show that the ZMPSTE 24 cleavage site on the precursor form of the lamin A tail domain orients itself in such a way as to facilitate cleavage during the maturation process. We confirm our simulated structures by comparing the thermodynamic properties of the ensemble structures to *in vitro* stability measurements. Using this combination of experimental and computational techniques, we compare the size, heterogeneity of size, thermodynamic stability of the Ig-fold, as well as the mechanisms of force-induced denaturation. Our data shows that the  $\Delta 50$  lamin A tail domain is more compact and displays less heterogeneity than the mature lamin A tail domain. Altogether these results suggest that the altered structure and stability of the tail domain can explain changed protein–protein and protein–DNA interactions and may represent an etiology of the disease. Also, this study provides the first molecular structure(s) of the lamin A tail domain, which is confirmed by thermodynamic tests in experiment.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

#### 1.1. Lamin A, laminopathies and HGPS

The nucleoskeleton of a cell, located at the inner face of the inner nuclear membrane, is composed mainly of intermediate filaments (IFs) made of either A-type and B-type lamin protein along with other accessory proteins (Wilson and Berk, 2010). A-type lamins, primarily lamin A and lamin C, are spliced from the *LMNA* gene. B-type lamins are encoded by *LMNB1* and *LMNB2*. Lamin A, B-type lamins are encoded by *LMNB1* and *LMNB2*. Lamin A,  $\Delta 50$  lamin A, HGPS  $\Delta 50$  mutant lamin A.

The nucleoskeleton of a cell, located at the inner face of the inner nuclear membrane, is composed mainly of intermediate filaments (IFs) made of either A-type and B-type lamin protein along with other accessory proteins (Wilson and Berk, 2010). A-type lamins, primarily lamin A and lamin C, are spliced from the *LMNA* gene. B-type lamins are encoded by *LMNB1* and *LMNB2*. Lamin A,  $\Delta 50$  lamin A, HGPS  $\Delta 50$  mutant lamin A.

**Abbreviations:** AA, amino acid; CD, circular dichroism; DLS, dynamic light scattering; DSC, differential scanning calorimetry; HGPS, Hutchinson–Gilford progeria syndrome; Ig, immuno-globulin; REMD, replica exchange method; SMD, steered molecular dynamics; WLC, worm-like chain; mwt LA, mature wild type lamin A;  $\Delta 50$  LA, HGPS  $\Delta 50$  mutant lamin A.

\* Corresponding authors. Addresses: Department of Biomedical Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA (K.N. Dahl); Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Mass. Ave., Room 1-235a&B, Cambridge, MA 02139 USA (M.J. Buehler).

E-mail addresses: [kndahl@andrew.cmu.edu](mailto:kndahl@andrew.cmu.edu) (K.N. Dahl), [mbuehler@MIT.EDU](mailto:mbuehler@MIT.EDU) (M.J. Buehler).

<sup>1</sup> These authors contributed equally to this work.

#### 1.2. Lamin A, molecular structure and the HGPS $\Delta 50$ mutation

Lamin A is a characteristic type V IF protein that contains a globular N-terminal head, a segmented coiled-coil  $\alpha$ -helical rod domain and a C-terminal tail containing an immunoglobulin (Ig)-fold (Herrmann et al., 2007). Lamin proteins are unique from other

IFs as they feature an exceptionally long C-terminal tail domain (Herrmann et al., 2007). The Ig-fold binds DNA and many other nuclear proteins (Zastrow et al., 2004). The C-terminus of the tail domain undergoes posttranslational processing, where the precursor form of lamin A is farnesylated, carboxymethylated, localized to the inner nuclear membrane (Coffinier et al., 2010) and then the last 18 amino acids are cleaved by an endoprotease ZMPSTE-24 to produce a mature wild-type lamin A (mwt LA) (Young et al., 2005). In HGPS, a single point mutation in the *LMNA* gene activates a cryptic splice site causing 50 amino acids encoded by exon 11 to be deleted, and the resulting mutant protein is called Δ50 lamin A (Δ50 LA) (De Sandre-Giovannoli et al., 2003). The deletion in Δ50 LA includes the ZMPSTE-24 cleavage site resulting in the retention of the C-terminal farnesylation, which is suggested to be responsible for the accumulation of Δ50 LA at the inner nuclear membrane. Similarly, the loss of the ZMPSTE-24 protease causes an accumulation of the precursor lamin A protein, prelamin A, at the inner nuclear membrane (Navarro et al., 2004; Taimen et al., 2009). However, the retained farnesylation cannot explain all of the molecular changes in HGPS. Recently, *in vitro* binding assays have shown differential binding of Δ50 LA to nuclear proteins and chromatin (Bruston et al., 2010). Two transgenic mice models containing an unfarnesylated Δ50 LA showed varied, but present, clinical pathology (Yang et al., 2011; Davies et al., 2010; Leuba et al., 1994). These results suggest that the loss of 50 amino acids from the lamin A tail may alter the protein more than simply retaining a farnesylation (Young et al., 2006).

### 1.3. Lamin A tail is an intrinsically disordered protein

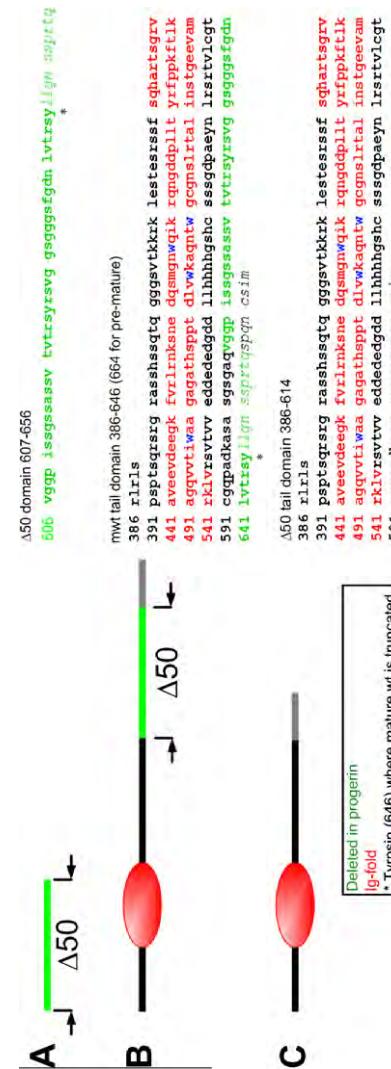
The tail domain of lamin A is mostly disordered and demonstrates the characteristic qualities of intrinsically disordered proteins (Rauscher and Pomes, 2010) including a promiscuity in protein binding (Schirmer and Foisner, 2007; Zastrow et al., 2004), tendency to aggregate (Linding et al., 2004) and a high proline and glycine content. It is difficult to predict how the removal of 50 amino acids in a region lacking secondary structure will affect the overall structure of the protein domain. Disordered regions are typically difficult to crystallize and take on many conformations, making traditional structural analyses such as X-ray crystallography and NMR challenging (Krimm et al., 2002; Zhang et al., 2000). Instead, here we use Replica Exchange Molecular Dynamics (REMD) simulations to determine an array of semi-stable structures based on ordered secondary structures (Sugita and Okamoto, 1999). REMD has been shown to be successfully utilized as a computational method used to improve the dynamic properties of conventional Molecular

Dynamics and Monte Carlo methods, aiming at obtaining global minimum free energy states of polymers by overcoming kinetic trapping without giving any specific initial structures (Liu et al., 2005; Snow et al., 2005; Sugita and Okamoto, 1999; Zhou et al., 2001). Nevertheless, this method can reduce accuracy when it is applied to extremely large systems (Snow et al., 2005). We focus on the amino acid sequence of the tail domains and apply massively parallelized computational power to study their intrinsically disordered structures (Liu et al., 2005) (see Section 4 and Supplementary Material for detail). When taken together, these structures represent the spectrum of conformations of the mwt LA and Δ50 LA tail domains. Then, we computationally and experimentally probe the protein stability to determine global structural changes in the tail domains. These structural changes likely impact binding, filament formation as well as protein stability. Specifically, we test the stability of the structured Ig-fold, which has previously been examined in other laminopathies (Krimm et al., 2002; Shumaker et al., 2005). However, an altered stability has only been observed for point mutations which occur in the Ig-fold. By combining data obtained from experiment and simulation, we extract information on changes in the stability of both normal and mutant proteins and develop a mechanistic model of higher-level nuclear structural and mechanical changes in diseased and normal states.

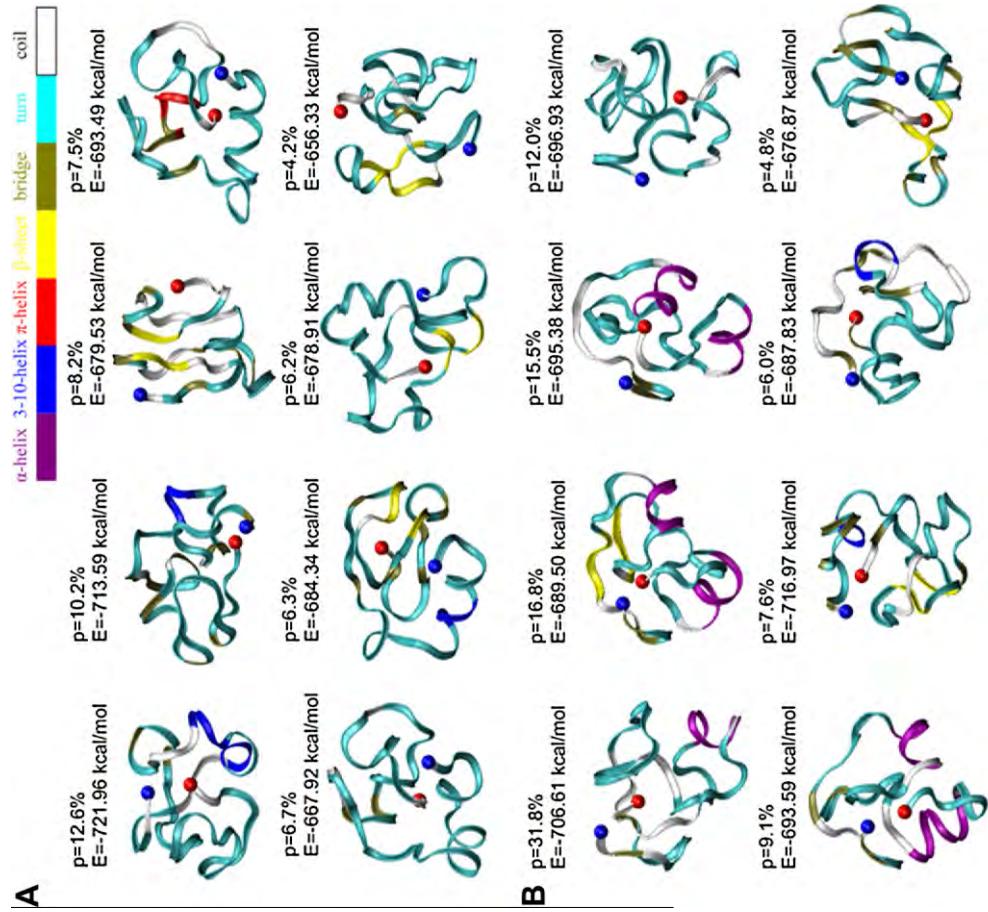
## 2. Results

### 2.1. Structure of the Δ50 region of lamin A

By applying REMD, we identify folded structures of different segments within the tail domain of LA. As shown in Fig. 1, the segments include AA607–656 (Δ50) for the 50 amino acids coded in exon 11 that are missing in the Δ50 LA mutation (sequence in Fig. 1A), AA428–646 for the mwt LA tail domain (sequence in Fig. 1B) and AA428–614\* for the Δ50 LA (sequence in Fig. 1C). Since the Δ50 region from exon 11 is lost in Δ50 LA, we examine the folded conformation of that region. To increase the sampling space and to examine the dependence on the initial geometry, we consider two extreme initial conformations; a straight conformation and an  $\alpha$ -helical conformation. This strategy has been effectively performed in earlier REMD studies applied to investigate protein structures (Liu et al., 2005). Likely structures of the Δ50 domain, that is, those associated with highest significance factors and lowest energies are listed in Fig. 2A and B. There are no dominant conformations for the structure based on the energy calculations of the different structures. The average end-to-end length of this segment is determined to be  $10 \pm 3.9 \text{ \AA}$  (Fig. S1C), obtained by statistical



**Fig. 1.** Schematic of the sequences of lamin tail domains investigated in this study. (A) The segment of amino acids deleted in the Δ50 LA tail. (B) The mwt LA tail ends at tyrosine 646 (\*), and the remaining residues (italics) are cleaved during posttranslational processing (these residues are included in our simulations). (C) Δ50 LA, also called progerin, has lost 50 amino acids (green) including the proteolytic site. Thus, Δ50 LA ends at a different 614, which we call 614\*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 2.** Conformations of the 50 amino acid segment of exon 11 with high significance factors. Simulations of protein folding with initial conditions of (A) straight and (B)  $\alpha$ -helix show similar final results, which are mostly disordered. Each amino acid is colored according to their secondary structure. The significance factor ( $p$ ) as well as the potential energy ( $E$ ) is marked for each conformation. The starting point (AA607) is in blue and the end point (AA656) is in red; (C) Secondary structure of the 50 amino acid segment shows little structure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

measurement of all the conformations generated. The flexibility of this segment is confirmed by comparing this end-to-end length with the contour length of 175 Å. These observations correlate with the experimental finding that the structure of this domain is amorphous and flexible (Kobayashi et al., 2000; Krimm et al., 2002). We confirm that two sets of initially differently folded structures converged to similar conformations characterized by the secondary structure distribution as shown in Fig. 2C. Fig. 2C clearly shows that the generated structures have a secondary structure indicative of intrinsic disorder, with >80% turns and coil. This is consistent with our experimental CD measurements (Fig. S2 and Table 1). A high glycine content (20%) in this region that may

**Table 1**  
Fit CD data (Whitmore and Wallace, 2004, 2008) of both tail domains shows a majority of disorder with  $\approx 25\%$   $\beta$ -sheet structure.

LA tail domain	Fraction of $\beta$ -sheet structure	Fraction of disorder
mwt	$0.27 \pm 0.05$	$0.66 \pm 0.05$
$\Delta 50$	$0.23 \pm 0.07$	$0.70 \pm 0.07$

be the source of the high flexibility of this segment. We observe a relatively high  $\alpha$ -helical character for  $\Delta 50$  segments that were initialized as a pure  $\alpha$ -helix (Fig. 2B). The  $\alpha$ -helix character is retained near the C-terminal of this  $\Delta 50$  domain around AA645 (8 AAs) and

623 (6 AAs), suggesting that they are more difficult to unfold rather than to maintain.

Another intriguing character of the  $\Delta 50$  region is that AA645–647 are always found on the outside of the structure with side-chains of the amino acids pointing outward (Fig. S1A and B in red). AA645–647 are the site of endoproteolytic cleavage by the protease, ZMPSTE-24, cleaving prelamin A to mwt LA (shown in Fig. 1B). Our computational results suggest that this “substrate” region faces outward possibly to facilitate interacting with the protease. This location may be critical to ensure the fast protease reaction kinetics of premature lamin A cleavage observed in cells (Liu et al., 2010).

## 2.2. Structures of tail domains of mwt LA and $\Delta 50$ LA

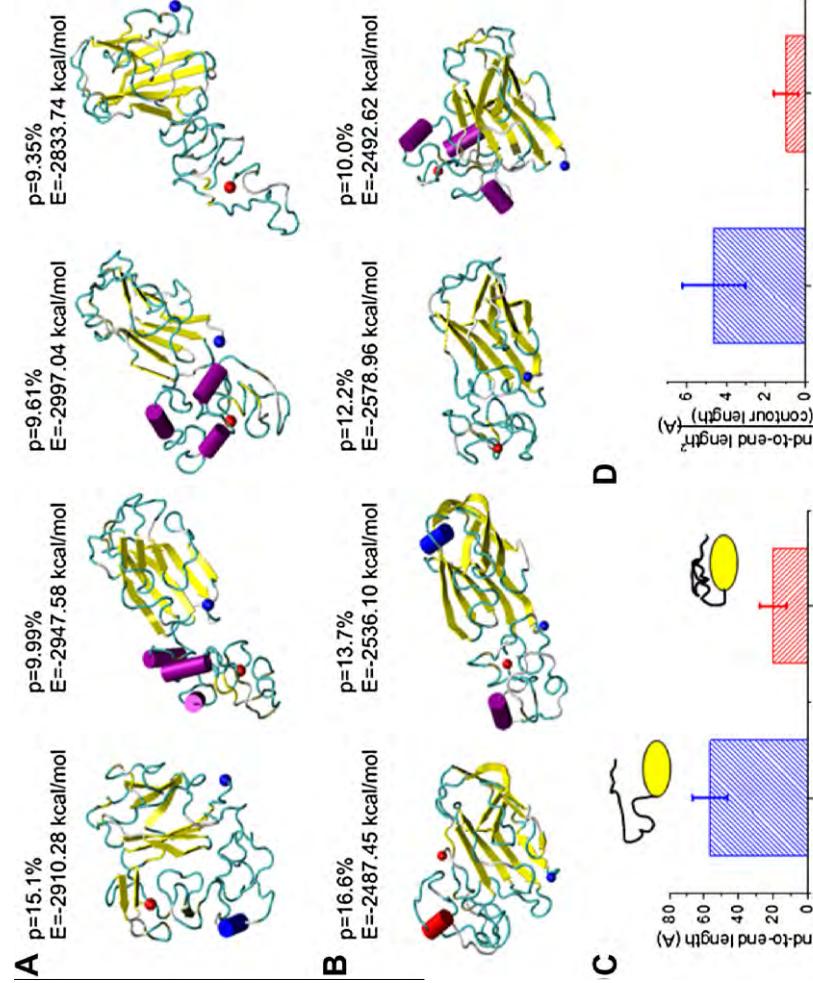
Experimentally, the tail domains of mwt LA and  $\Delta 50$  LA show a  $\beta$ -sheet structure with random coil but no  $\alpha$ -helix (Fig. S2, Table 1). After purification, the proteins are estimated to be  $\approx 4.0 \pm 0.3$  nm in size, as measured by DLS (Fig. S3). Using the same REMD technique as used above, we predict conformations of the full tail domains of mwt LA and  $\Delta 50$  LA. As expected for a mostly intrinsically disordered protein, our simulations yield an ensemble of conformations for the tail domains of mwt LA and  $\Delta 50$  LA. We present the eight conformations that are most frequently visited by the replica with the lowest temperature in the order of decreasing significance (Fig. 3A and B). Initially, the structure of the s-type Ig-fold of the tail domain is modeled from the starting conformation available from the protein database (PDB with ID 1IVT (Krimm et al., 2002)), and the remaining tail domain is given a straight conformation.

623 (6 AAs), suggesting that they are more difficult to unfold rather than to maintain. Another intriguing character of the  $\Delta 50$  region is that AA645–647 are always found on the outside of the structure with side-chains of the amino acids pointing outward (Fig. S1A and B in red). AA645–647 are the site of endoproteolytic cleavage by the protease, ZMPSTE-24, cleaving prelamin A to mwt LA (shown in Fig. 1B). Our computational results suggest that this “substrate” region faces outward possibly to facilitate interacting with the protease. This location may be critical to ensure the fast protease reaction kinetics of premature lamin A cleavage observed in cells (Liu et al., 2010).

We find that the folded structures of  $\Delta 50$  LA tail (Fig. 3B) are more compact than structures of mwt LA tail (Fig. 3A). From these simulation results, we quantify the end-to-end length ( $r_{ee}$ ) of this disordered domain (AA545–646 for mwt LA and AA545–614 for  $\Delta 50$  LA in Fig. 3C). The average  $r_{ee}$  for mwt LA tail was measured to be  $r_{ee} = 56.5 \pm 10.2$  Å, which is much larger than that of the  $\Delta 50$  LA tail ( $r_{ee} = 20.4 \pm 7.6$  Å). Some of the reduction in the  $r_{ee}$  in the  $\Delta 50$  LA tail, as compared to mwt LA tail counterpart, can be accounted for by the reduced contour length from the net loss of 32 amino acids. However, the contour length is only reduced by 30%. much less than their difference in  $r_{ee}$ . The persistence length of an isolated polypeptide chain is  $\approx 4$  Å, which is much smaller than the loss of contour length (357 Å for 102 AAs of mwt LA versus 245 Å for 70 AAs of  $\Delta 50$  LA). As such, we can use the Worm-Like-Chain (WLC) model for the disordered regions (in the limit of  $l \gg \xi_p$ ) (Leuba et al., 1994; Marko and Siggia, 1995)

$$\xi_p = \frac{\langle r_{ee}^2 \rangle}{2l} \quad (1)$$

to estimate the persistence lengths of the mwt LA tail and the  $\Delta 50$  LA tail. We obtain that the persistence length of the  $\Delta 50$  LA tail ( $1.0 \pm 0.6$  Å) is 78% shorter than that of the mwt LA tail ( $4.6 \pm 1.6$  Å), as shown in Fig. 3D. This result indicates that the



**Fig. 3.** Conformations and structures of tail domains for normal and disease cases. Conformations with high significance factors of the (A) mwt LA tail domain AA428–646 and (B)  $\Delta 50$  LA tail domain AA428–614\* show different structures. The beginning amino acid is marked with a blue dot and the C-terminus is marked with a red dot. (C) The end-to-end length from the end of the Ig-fold domain (AA545) to the tail end (AA646 for mwt LA and AA614\* for  $\Delta 50$  LA) of all the conformations of the mwt LA and  $\Delta 50$  LA tail domains, respectively. (D) Normalizing the square of the end-to-end length by the contour length reflects the persistence length of the peptide. We find that the  $\Delta 50$  LA tail domain is significantly more compact, even when considering the loss of amino acids in the truncation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 2**  
Hydrodynamic radius of the tail domain measured by DLS and simulation.

LA tail domain	DLS experiment (nm)	Simulation (nm)
mwt	4.5 ± 1.3	5.0 ± 1.4
Δ50	4.0 ± 0.3	3.2 ± 0.3

tail domain of Δ50 LA is more able to adopt a greater variety of conformations. We calculate the effective radius of the molecule, and find that the diameter of Δ50 LA tail closely agrees with the experimental hydrodynamic radius from DLS (Table 2; Fig. S4 to compare with Fig. S3). Interestingly, the deviation of the protein sizes, both by experiment and simulation, are the most striking features (Table 2); and the mwt LA has a more heterogeneous distribution of protein sizes.

We quantitatively measure the overall structural composition of the mwt LA and Δ50 LA for the tail domain, and we observe that >70% of the amino acids feature turn or coil structures, confirming that this domain is mostly disordered (Fig. S5A). We count the secondary structures of each amino acid and plot the ratio (Fig. S5B and C). Some positions in the mwt LA tail are seen to form  $\alpha$ -helices including 562 (24 AAs), 584 (8 AAs) and 620 (20 AAs). In the Δ50 LA tail,  $\alpha$ -helices are similar if not slightly longer for the mid-segment but are missing the third segment (562 (24 AAs) and 582 (12 AAs)). Compared with random coils, a peptide with a helix structure has a much longer persistence length ( $\approx 100$  Å since the backbone is stabilized by hydrogen bonds), which can explain why the mwt LA tail, with net 16 more AAs in helices, has a longer persistence length.

### 2.3. The thermal denaturation of the mature lamin A and Δ50 LA tail domains

To experimentally measure the stability of the LA tails we utilize the presence of four hydrophobic tryptophan residues in the Ig-fold (Fig. 1) as a way to monitor protein thermal unfolding (Efteinik, 1994). We monitor the intensity of protein solutions at 342 nm, the peak wavelength of the emission spectra, as a function of temperature. The intensity is then normalized assuming a fully folded state at low temperature 37 °C and a fully unfolded state at high temperature 95 °C (see Section 4 and Fig. S6). The resulting unfolding curve as a function of temperature is shown in Fig. 4A. Measuring the thermal stability gives us insight into the tail domains' thermal equilibrium ( $k_B T$ ). Using this analysis, we find that the mwt LA tail shows little sample-to-sample variation and we determine a transitional melting temperature ( $T_{M,mwt}$ ) of  $63.2 \pm 0.09$  °C at which 50% of the protein is unfolded. The Δ50

**Table 3**  
Experimentally measured  $T_M$  of lamin A tails from thermal denaturation and tryptophan fluorescence.

LA tail domain	$T_M$ (°C)	$r_i^2$	$\sigma_i$ (°C)	$T_M$ (°C)	$\sigma_{T_M}$ (°C)
mwt	63.13, 63.24	0.96, 0.94	0.070, 0.055	63.2	0.09
Δ50	65.80, 65.82	0.32, 0.88	0.215, 0.085	65.8	0.23

LA tail domain shows an increased sample-to-sample variation and shows a  $T_{M,\Delta 50}$  of  $65.8 \pm 0.23$  °C (Fig. 4B, Table 3).

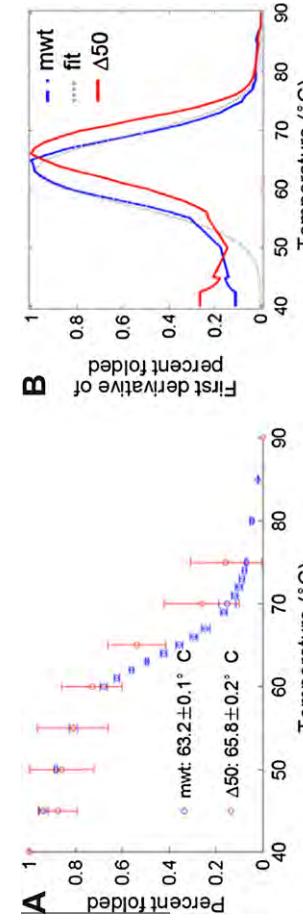
We confirm the melting temperature of the Δ50 LA tail by examining changes in the secondary structure using CD with increasing temperature as well as by differential scanning calorimetry (DSC). We perform these analyses on Δ50 LA tail domain since the protein is less susceptible to aggregation and the baseline protein state was more regular (see Supplementary Material). The change in CD signal at 230 nm with increasing temperature normalized to room temperature and denatured points shows a characteristic denaturation when fit to a van't Hoff model.

$$\ln(K_{eq}) = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \quad \text{where the } K_{eq} = (\text{fraction unfolded})/(\text{fraction folded}). \quad (2)$$

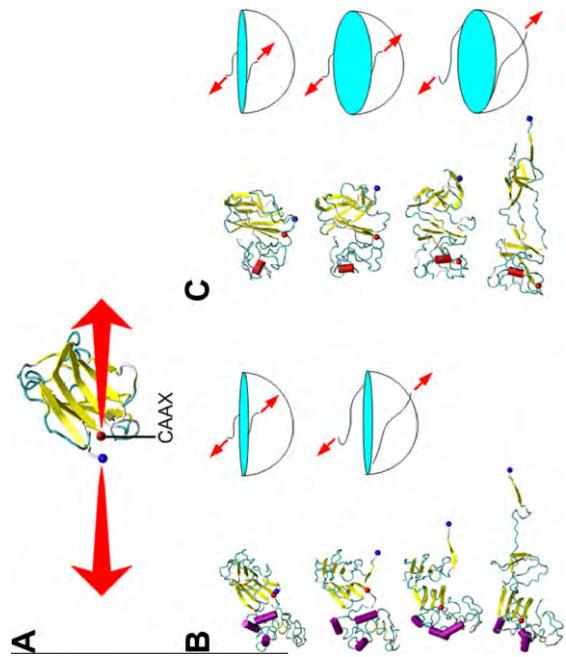
From this fit (Fig. S6A), we observed a characteristic  $T_M$  of 63 °C based on secondary structure. We confirm this transition temperature using DSC, and observed a  $T_M$  of 63 °C (Fig. S6B). These values are slightly below the  $T_M$  proposed earlier, the similar range of temperatures for denaturation of secondary and tertiary structure suggests robustness of the thermal stability measures. However, changes in secondary structure are limited by low signal to noise ratio (small amounts of  $\beta$ -sheet compared to random coil).

### 2.4. Simulating the thermodynamic denaturation of tail domains under loading conditions

To complement the experimental data of protein stability described in the previous section, we use simulation to measure the mechanical stability of the mwt and Δ50 LA tail domains by unfolding the protein at 27 °C and monitoring the conformation of the region containing the Ig-fold. We focus on the Ig-fold because: (a) the Ig-fold is the most ordered and probably most stabilized motif representing the limiting factor of mechanical strength; (b) the Ig-fold interacts with laminin binding proteins and DNA (Stierle et al., 2003; Zastrow et al., 2004); and (c) we can directly compare denaturation with experimental results based on the tryptophan residues in the Ig-fold. We apply the steered molecular dynamics (SMD) protocol to stretch the two regions flanking the Ig-fold domain (AA428 and 545) within the full tail domain to



**Fig. 4.** Thermal denaturation of Δ50 LA and mwt LA protein by measuring tryptophan fluorescence. (A) Fluorescence of maximum tryptophan signal are normalized to signals corresponding to 100% folded (measured at 37 °C) and 0% folded (at 95 °C). The averages of two independent runs are shown of each protein. (B) To find the transition melting temperature, we take the first derivative of the signal. Interpolation, smoothing and fit to a Gaussian curve (see Fig. S6) allow us to determine transition temperatures with confidence intervals (Table 3).

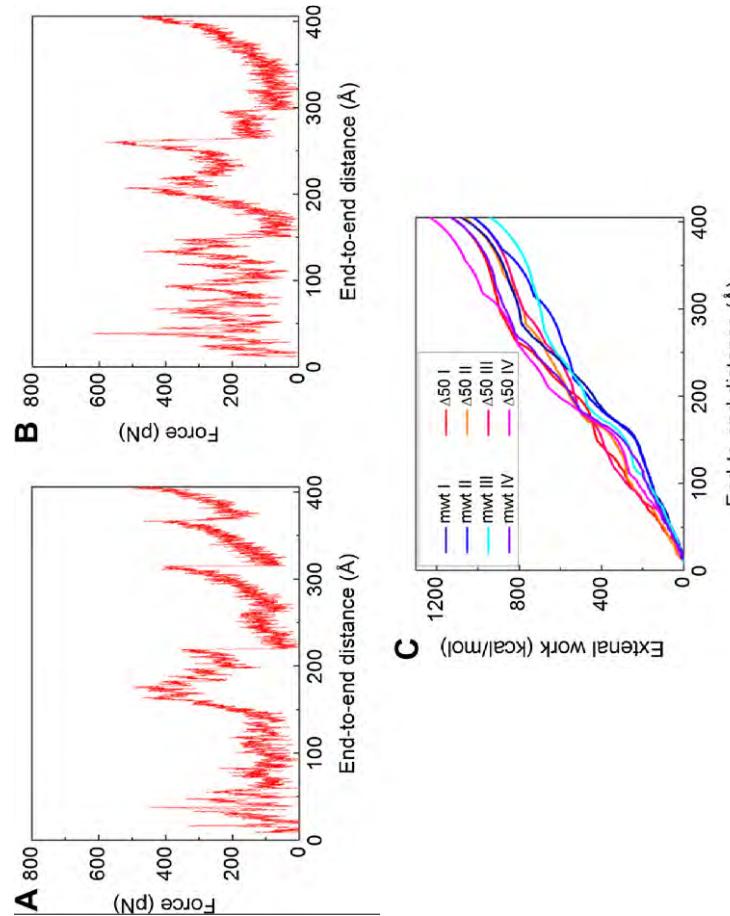


**Fig. 5.** Results of SMD simulation of Ig-fold domain unfolding. (A) The  $\beta$ -sheet structure of the Ig-fold domain and boundary conditions used to apply the mechanical load. While the force was applied to the Ig-fold, the other amino acids were in their pseudo-equilibrium states. (B) Snapshots of a mwt LA tail domain during the unfolding process, taken at extensions of 0, 15, 38 and 75 Å, respectively show one  $\beta$ -strand at a time (yellow) unfolding from a central structure. (C) Snapshots of a  $\Delta 50$  LA tail domain during the unfolding process, taken at extensions of 0, 15, 38 and 75 Å, respectively. The images show an intermediate state with no “outlying”  $\beta$ -strands. The cartoons in panels B and C schematically describe the different deformation process of the mwt and  $\Delta 50$  LA tail domains, respectively. Note the large blue deformation pocket in panel C is larger for  $\Delta 50$  LA tail domain, than that for mwt LA tail domain and present for larger deformations (and representing the intermediate state).

unfold the structure and record the force–extension curve during the stretching process (Fig. 5A).

We find that the unfolding processes of the two protein segments reveal different deformation mechanisms under loading. The unfolding of the Ig-fold in the mwt LA tail shows a stepwise unwinding of the  $\beta$ -strands from the one end to the other end (Fig. 5B). However, the unfolding of the Ig-fold of the  $\Delta 50$  LA tail domain reveals a two-state process. In the first state, the Ig-fold maintained secondary  $\beta$ -sheet structure but shows a tertiary transition, similar as opening two halves of a sandwich along an edge. This unfolding pattern opens up beta-sheet pocket. In the second stage, beta-strands are peeled away from the central structure in a similar fashion as mwt LA tail (Fig. 5C). The two-state unfolding process of the Ig-fold domain appears to be unique to the  $\Delta 50$  LA tail, which could qualitatively describe by a pocket-unwinding model as illustrated in Fig. 5C. This result shows that the Ig-fold of the  $\Delta 50$  LA tail transforms before unfolding under loading. These results suggest altered interactions within the Ig-fold and that the mutation makes unfolding more variable. Thus, the two-state unfolding complicates a normally continuous process and may allow for more stable intermediate states.

We repeat the loading process for other possible structures with high significance factor (structures in Fig. 3A and B) and collect all force–extension curves as shown in Fig. 6A and B from the initial conformation until the end-to-end distance equaled the contour length ( $\sim 403$  Å). We model force–extension curves as worm-like-chain (WLC) models stretching at the end (Marko and Siggia, 1995). Initially, there were 4–5 transition peaks, each corresponding to an unfolding event of a  $\beta$ -strand, requiring the unfolding force of  $357.8 \pm 94.2$  pN. After each unfolding event, the force drops and shifts rightward because the contour length has increased (Sotomayor and Schulten, 2007). We integrate the force–extension curves (Fig. 6C) and find that the external work to fully unfold the



**Fig. 6.** Force–extension curves of the unfolding of Ig-fold domains. (A) A force–extension curve as recorded during the unfolding process of the mwt LA Ig-fold as shown in Fig. 5B. (B) A force–extension curve as recorded during the unfolding process of the  $\Delta 50$  Ig-fold as shown in Fig. 5C. (C) The external work on the Ig-fold domain during the unfolding process as a function of the end-to-end distance, which is computed by integrating of the force–extension curves obtained by unfolding each of the structures shown in Fig. 3A and B.

$\Delta 50$  LA Ig-fold was  $1113.8 \pm 75.1$  and  $1057.3 \pm 70.2$  kcal/mol for the mw<sup>t</sup> LA Ig-fold. To fully unfold the Ig-fold of the  $\Delta 50$  LA tail domain, 56.5 kcal/mol more work is required than in the case of the mw<sup>t</sup> LA tail domain. We believe that this extra energy is required because the  $\Delta 50$  LA tail is more compact, which stabilizes the Ig-fold and results in an increase of the unfolding energy. From Fig. 6C, the energy difference at 75 Å extension, corresponding to the first-stage transition shown in Fig. 5C, is 65 kcal/mol, suggesting that the additional energy to fully unfold the Ig-fold may primarily be required for the first-state, hinge-like transition.

We now qualitatively compare our *in vitro* result with the *in silico* results and find that the  $\Delta 50$  LA tail is more stable in both experiment and simulation. Experimentally, the thermal transition mid-points are  $T_{M,\Delta 50} = 65.8 \pm 0.23$  °C for the  $\Delta 50$  LA tail domain and  $T_{M,mw^t} = 63.2 \pm 0.09$  °C for the mw<sup>t</sup> LA tail domain, indicating that the  $\Delta 50$  LA tail domain is more difficult to unfold. This experimental result agrees with the *in silico* result that additional energy was required to fully unfold the  $\Delta 50$  LA tail domain structure. According to a calculation based on the Bell-Evans model (Bell, 1978; Evans and Ritchie, 1997; Rief et al., 1998), we arrive at  $\Delta W_{exp} = 37.0 \pm 7.0$  kcal/mol ( $62 \pm 12$   $k_B T$  for  $T = 300$  K). This experimental result corresponds to the additional energy of  $\Delta W_{sim} = 69.8 \pm 17.4$  kcal/mol ( $117 \pm 29$   $k_B T$  for  $T = 300$  K) needed to unfold 50% of the  $\Delta 50$  LA tail more than the mw<sup>t</sup> LA tail in our computational studies. We find that the computational results are on the same order of magnitude as the experimental result. The difference between the experiment and simulation results could be caused by the intrinsic difference of the two methods. In the SMD simulation approach, the force is applied only at the two ends of the strand with a finite loading rate, the structure unfolded along the prescribed direction and the force required to overcome the energy barriers is applied only in this one prescribed direction. In temperature mediated changes in tryptophan fluorometry, the unfolded form has more degrees of freedom and the rate of temperature increase is also intrinsically different from the loading rate used in our simulation. Also, experiments could possibly include simultaneous kinetic aggregation, which was not included in the simulation (see Section 4 for detail).

### 3. Discussion

Using a combination of theoretical and experimental methods, we have studied atomic conformations and nanomechanical stabilities of the tail domain of the mw<sup>t</sup> LA tail and a mutant associated with HGPS ( $\Delta 50$  LA tail). Our results demonstrate that the  $\Delta 50$  LA tail is more compact than the mw<sup>t</sup> LA tail domain. Most importantly the unfolding of the  $\Delta 50$  LA tail domain requires a higher energy barrier to overcome unfolding, likely due to transitioning through a pseudo-stable intermediate state (Figs. 3–6 and Table 1). This phenomenon is in agreement with experimental measurements using both tryptophan fluorometry and circular dichroism methods (Figs. 4 and S7).

These findings highlight the need to keep the protein geometry in mind when trying to understand the molecular mechanisms of the disease. Previously, experimental works of wild type and HGPS cells have shown a stiffening of cell nuclei from HGPS patients and resistance to mechanical pressure, and alteration in the mechanical properties is due to the presence and overaccumulation of the  $\Delta 50$  LA protein and associated filaments at the inner nuclear membrane (Dahl et al., 2004; Goldman et al., 2004). The etiology of HGPS has long been thought to be a function of the retained farnesyl group alone (Capell et al., 2008; Leuba et al., 1994), but our results suggest that changes in protein stability may be contributing to the disease pathology. This finding is consistent with recent studies which show that the presence of unfarnesylated prelamin A and  $\Delta 50$  LA still have cellular and organism effects whose mechanisms

are not yet entirely clear (Yang et al., 2011; Davies et al., 2010; Verstraeten et al., 2008).

Furthermore, our experimental and computational results quantitatively and consistently show that the  $\Delta 50$  LA tail is a more stable structure (by approximately 37–70 kcal/mol or 62–117  $k_B T$ ), as well as a more compact molecule. This suggests an explanation for the increased compactness of *in vitro* reconstituted filaments of full length  $\Delta 50$  LA than mw<sup>t</sup> LA (Taimen et al., 2009). This compact filament structure, likely caused by a smaller and more stable  $\Delta 50$  LA tail, may also partly be responsible for reduced turnover of proteins from the nucleoskeleton (Dahl et al., 2004; Goldman et al., 2004) and is thus much more than just a farnesylation-related issue. We also observe an increased propensity for  $\Delta 50$  LA tail domain to form dimers, suggesting an increase in dimer–dimer interaction strength and as a result, a reduced capacity for the  $\Delta 50$  LA tail domain to be reorganized by thermal fluctuations (moderated temperatures or moderate force), consisted with the extra energy and higher temperature that is required to unfold the  $\Delta 50$  LA tail domain seen here. Filaments that include the more stable  $\Delta 50$  LA may more readily aggregate, show reduced turnover, and due to their stronger binding appear mechanically more rigid and therefore “brittle”, thereby directly explaining earlier experimental findings (Dahl et al., 2004, 2006; Delbarre et al., 2006) based on the geometric and mechanistic insight developed in this study. Our results may also explain a recent experimental report that showed that the proliferation of HGPS cells is significantly impaired (Verstraeten et al., 2008).

From a mechanistic viewpoint we have shown that the mw<sup>t</sup> LA tail unfolds by individual beta-strands being pulled away from the central structure in a sequential process. In contrast, the  $\Delta 50$  LA tail domain shows a two-state transition process with a tertiary opening of the Ig-fold into two halves, which we termed a “beta-sheet pocket”. Within this pocket there may exist other “hidden” binding sites similar to the internal structure of the Ig-folds of titin (Marszałek et al., 1999) or fibronectin (Klotzsch et al., 2009) that may alter the affinity to binding partners. The two stage transition process may also alter the unfolding process of the assembled filament, and we hypothesize that this mechanism could imply that the unfolding of  $\Delta 50$  LA lamin filament features a greater structural variability, while the process in case of wild-type lamin filament is smoother. This could explain differential binding of  $\Delta 50$  LA to proteins and DNA (Bruston et al., 2010). We suggest that there may be other differential protein–protein binding associated with  $\Delta 50$  LA, but this binding may be energy dependent, such as with force or temperature.

The “structure” of intrinsically disordered proteins is described as an ensemble of conformations which may represent conformations allowing for protein–protein binding to different binding partners and have been shown to be necessary for specific mechanical functions of biological materials including spider silk and tailin rods (del Rio et al., 2009; Keten et al., 2010). Given Lamin A’s proposed role in mechanotransduction, and the multiple binding partners of the tail domain, further study of the conformations of the tail domain relating to mechanical function should be explored in future work (Wang et al., 2009). Here, we have demonstrated the use of an integrated experimental–computational approach to determine a group of structures to gain molecular-level insight into pathological mechanisms.

### 4. Materials and methods

#### 4.1. Protein concentration, purity and size characterization

Protein concentration was measured using Coomassie Plus (Bradford) Assay (Pierce). Purified protein size and purity was

determined using 14% SDS-PAGE, according to standard protocols. Dynamic light scattering (Malvern Zeta Sizer Nano) was used to measure protein hydrodynamic radius and possible aggregation state.

#### 4.2. Protein expression and purification

Lamin A tail domains were created using PCR from plasmids kindly provided by Misteli at NCI/NIH (Dahl et al., 2004) to create tail fragments of mw<sub>t</sub> LA (386–646; Fig. 1B) and  $\Delta$ 50 LA (Fig. 1C) using primers. For the mw<sub>t</sub> LA tail domain with encoded EcoR1, Xba1 and stop codon we used: 5'-GAA TTC CCT ACC CCT GTC CCC CA-3' and 5'-CTC GAG TTA GTA GCA GCG GGT GAC CAG-3' and for  $\Delta$ 50 LA tail domain with EcoR1, Xba1 and stop codon we used: 5'-GAA TTC CTA CGC CTC TCC CCC AGC CC-3' and 5'-CTC GAG TTA CAT GAT GCT GCA GTT CTC GGG GCT-3'. PCR products were ligated into pGEMTeZ vectors, amplified in DH5 $\alpha$ , and subcloned into the pGST parallel vector for recombinant expression (Sheffield et al., 1999). GST-tagged lamin A tails were transformed into BL21-CodonPlus competent *Escherichia coli* (Agilent Technologies), grown 2–3 h at 37 °C, and induced by IPTG at OD600 of 1.8 (mw<sub>t</sub> LA: 400  $\mu$ M IPTG for 3 h;  $\Delta$ 50 = 800  $\mu$ M IPTG for 2 h).  $\Delta$ 50 LA tail was produced more easily and was less susceptible to degradation, possibly suggesting resistance to protease activity (Parsell and Sauer, 1989). Cells were harvested by centrifugation at 6000  $\times$  g for 10 min, lysed with B-PER reagent (Pierce), diluted by 5  $\times$  in binding/wash buffer: 125 mM Tris, 500 mM NaCl, pH 8.0 and 100  $\mu$ M PMSF and added to glutathione magnetic beads overnight at 4 °C (Pierce). Beads were washed with 20 volumes with binding/wash buffer and exposed to protease for 5 h at 30 °C according to manufacturer's instructions (Promega Madison, WI). Cleaved protein was collected in the supernatant after centrifugation (1000  $\times$  g for 5 min) and exposed to glutathione agarose beads (Pierce) pre-equilibrated in protease buffer to remove excess GST overnight at 4 °C. Proteins were ultracentrifuged at 100 000  $\times$  g for 30 min at 4 °C to remove aggregates.

#### 4.5. Extra energy needed to unfold Ig fold of $\Delta$ 50 LA tail

According to the Bell-Evans model (Bell, 1978; Evans and Ritchie, 1997; Rief et al., 1998), the unfolding probability of a structure stabilized by molecular bonds is proportional to  $\exp(-E_b/k_B T)$  where  $E_b$  is the energy barrier height and  $k_B T$  is the thermal energy. Thereby, for  $\Delta$ 50 LA at 50% unfolding in simulation, we have

$$\exp \left[ \frac{-(E_{\Delta 50} - W_{M,\Delta 50})}{k_B T_0} \right] = \exp \left[ \frac{-E_{\Delta 50}}{k_B T_{M,\Delta 50}} \right], \quad (3)$$

where  $\langle W_{M,\Delta 50} \rangle = 551.5 \pm 40.1$  kcal/mol is the averaged external energy applied to unfold 50% of the  $\Delta$ 50 LA as the end-to-end length equals to half of the fully unfolded length (Fig. 6C) and  $T_0 = 27$  °C is the room temperature. Eq. (3) means the increasing temperature has a same effect at the intermediate unfolding stage as the external energy as shown in Fig. S7, and from this equation we obtain  $E_{\Delta 50} = 481.7$  kcal/mol. In experiments:

$$\exp \left[ \frac{-E_{\Delta 50}}{T_{M,\Delta 50}} \right] = \exp \left[ \frac{-E_{mw\tau}}{T_{M,mw\tau}} \right], \quad (4)$$

And since both of the two structures are at the intermediate unfolding stage with a same unfolding probability. We have similar conclusion for the simulations as

$$\exp \left[ \frac{-(E_{\Delta 50} - W_{M,\Delta 50})}{k_B T_0} \right] = \exp \left[ \frac{-(E_{mw\tau} - W_{M,mw\tau})}{k_B T_0} \right]. \quad (5)$$

By solving Eqs. (4) and (5), we have  $\Delta W_{exp} = E_{\Delta 50}/(T_{M,\Delta 50} - T_{M,mw\tau})/T_{M,\Delta 50} = 37.0 \pm 7.0$  kcal/mol ( $= 62 \pm 12$   $k_B T$  for  $T = 300$  K), which corresponds to the extra energy needed to unfold 50% of the  $\Delta$ 50 LA more than the mw<sub>t</sub> LA. Since  $\langle W_{M,mw\tau} \rangle = 481.7 \pm 46.1$  kcal/mol is obtained by simulation from Fig. 6C, we obtain the extra energy term from simulation as  $\Delta W_{sim} = \langle W_{M,\Delta 50} - W_{M,mw\tau} \rangle = 69.8 \pm 17.4$  kcal/mol ( $= 117 \pm 29$   $k_B T$  for  $T = 300$  K), which is in the same order of magnitude as the experimental result.

#### Acknowledgments

Z. Qin and M.J. Buehler acknowledge support from AFOSR and DOD-PECASE. A. Kalinowski and K.N. Dahl acknowledge R. Tilton, T. Przybycien (Chemical Engineering), B. Armitage (Chemistry) and G. Rule (Biology) at Carnegie Mellon University for use of their shared equipment, as well as N. Kolluri, S. Chapman, P. Silberman and J. Wert for their assistances in experiments. Funding from the Progeria Research Foundation (to KND) and NIH (NRSA F30AG030905 from NIA to AK) is also acknowledged. We acknowledge helpful discussions with Tom Misteli.

#### 4.6. Tryptophan fluorometry

For tryptophan fluorescence tail domains were diluted to 0.1 mg/mL in 50 mM HEPES, 0.5 mM EDTA, pH 7.0. Sample was excited at 295 nm in a 1 cm pathlength polished cell in a Fluorolog-3 (Jobin Yvon-spx, Instruments S.A., Inc., Edison, NJ) and the emission at 342 nm was monitored as the temperature was increased from 40 to 90 °C in 1 °C increments,  $\approx$  every 1 min. The temperature was manually controlled using a water bath (Neslab RTE-210, ThermoScientific) attached to a custom built sample holder, allowing for temperature regulation.

#### 4.7. Molecular simulation approach and analysis

Replica Exchange Molecular Dynamics (REMD) and Steered Molecular Dynamics (SMD) simulations are carried out using the CHARMM19 all-atom energy function with an effective Gaussian model for the water solvent (additional details about the simulation approach and analysis methods are provided in Supplementary methods part). For the mechanical analysis the  $\alpha$ -carbon atoms at the two ends of the Ig-fold are connected to SMD springs (spring stiffness 10 kcal/ $\text{\AA}^2$ mol) and they move against each other at a constant relative speed of 0.01  $\text{\AA}/\text{ps}$  (Qin and Buehler, 2010; Qin et al., 2009). The applied pulling force is recorded versus its displacement. We use Visual Molecular Dynamics (VMD) (Humphrey et al., 1996) for visualization of protein structures (Figs. 2, 3 and 5). We assess secondary structures by analyzing both the hydrogen bond patterns and the backbone geometry with the STRIDE algorithm (Frishman and Argos, 1995).

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2011.05.015.

#### References

- Bell, G.I., 1978. Models for the specific adhesion of cells to cells. *Science* 200, 618–627.
- Bruston, F., Delbarre, E., Ostlund, C., Worman, H.J., Buendia, B., et al., 2010. Loss of a DNA binding site within the tail of prelaminin A contributes to altered heterochromatin anchorage by progerin. *FEBS Lett.* 584, 2999–3004.
- Capell, B.C., Olive, M., Erdos, M.R., Cao, K., Faddah, D.A., et al., 2008. A farnesytransferase inhibitor prevents both the onset and late progression of cardiovascular disease in a progeria mouse model. *Proc. Natl. Acad. Sci. USA* 105, 15902–15907.
- Coffinier, C., Jung, H.J., Li, Z., Nobumori, C., Yun, U.J., et al., 2010. Direct synthesis of lamin A, bypassing prelamin a processing, causes misshapen nuclei in fibroblasts but no detectable pathology in mice. *J. Biol. Chem.* 285, 20818–20826.

Dahl, K.N., Kahn, S.M., Wilson, K.L., Discher, D.E., 2004. The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. *J. Cell Sci.* 117, 4779–4786.

Dahl, K.N., Scaffidi, P., Islam, M.F., Yodh, A.G., Wilson, K.L., et al., 2006. Distinct structural and mechanical properties of the nuclear lamina in Hutchinson–Gillford progeria syndrome. *Proc. Natl. Acad. Sci. USA* 103, 10271–10276.

Davies, B.S., Barnes, 2nd, R.H., Tu, Y., Ren, S., Andres, D.A., et al., 2010. An accumulation of non-farnesylated prelamin A causes cardiomyopathy but not progeria. *Hum. Mol. Genet.* 19, 2682–2694.

De Sandre-Giovannoli, A., Bernard, R., Cau, P., Navarro, C., Courvalin, J.C., et al., 2003. Lamin A truncation in Hutchinson–Gillford progeria. *Science* 300, 2055.

Del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J.M., Sheetz, M.P., 2009. Stretching single talin rod molecules activates vinculin binding. *Science* 323, 638–641.

Delbarre, E., Tranier, M., Copey-Moisan, M., Gaillard, C., Courvalin, J.C., et al., 2006. The truncated prelamin A in Hutchinson–Gillford progeria syndrome alters structure of A-type and B-type lamin homopolymers. *Hum. Mol. Genet.* 15, 1113–1122.

Eftink, M.R., 1994. The use of fluorescence methods to monitor unfolding transitions in proteins. *Biophys. J.* 66, 482–501.

Evans, E., Ritchie, K., 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72, 1541–1555.

Fong, L.G., Ng, J.K., Lammerding, J., Vickers, T.A., Meta, M., et al., 2006. Prelamin A accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson–Gillford progeria syndrome. *Proc. Natl. Acad. Sci. USA* 101, 8953–8956.

Herrmann, H., Bar, H., Kreplak, L., Strelkov, S.V., Aeby, U., 2007. Intermediate filaments, from cell architecture to nanomechanics. *Nat. Rev. Mol. Cell Biol.* 8, 562–573.

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14 (33–8), 27–28.

Keten, S., Xu, Z.P., Ihle, B., Buehler, M.J., 2010. Nanoconfinement controls stiffness, strength and mechanical toughness of beta-sheet crystals in silk. *Nat. Mat.* 9, 359–367.

Klotzsch, E., Smith, M.L., Kubow, K.E., Muntwyler, S., Little, W.C., et al., 2009. Fibronectin forms the most extensible biological fibers displaying switchable force-exposed cryptic binding sites. *Proc. Natl. Acad. Sci. USA* 106, 18267–18272.

Kobayashi, N., Kigawa, T., Koshiba, S., Inoue, M., and Yokoyama, S. Solution structure of immunoglobulin-like domain of mouse nuclear lamin. RCSB Protein Database. doi:10.1221/pdb.1ufg/pdb.

Krann, I., Ostlund, C., Gilquin, B., Couprie, J., Hossenlopp, P., et al., 2002. The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy. *Structure* 10, 811–823.

Lammerding, J., Fong, L.G., Ji, J.W., Reue, K., Stewart, C.L., et al., 2006. Lamins A and C but not lamin B1 regulate nuclear mechanics. *J. Biol. Chem.* 281, 25768–25780.

Leuba, S.H., Yang, G., Robert, C., Samoni, B., van Hoide, K., et al., 1994. Three-dimensional structure of extended chromatin fibers as revealed by tapping-mode scanning force microscopy. *Proc. Natl. Acad. Sci. USA* 91, 11621–11625.

Linding, R., Schymkowitz, J., Rousseau, F., Diella, F., Serrano, I., 2004. A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J. Mol. Biol.* 342, 345–353.

Liu, J., Kim, B., Berne, B.J., 2005. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. USA* 102, 13749–13754.

Liu, Q., Kim, D.I., Syme, J., LuValle, P., Burke, B., et al., 2010. Dynamics of lamin-A processing following precursor accumulation. *PLOS One* 5, e10874.

Marko, J.F., Sigga, E.D., 1995. Stretching DNA. *Macromolecules* 28, 8759–8770.

Marszałek, P.E., Lu, H., Li, H., Carrion-Vazquez, M., Oberhauser, A.F., et al., 1999. Mechanical unfolding intermediates in titin modules. *Nature* 402, 100–103.

Navarro, C.I., De Sandre-Giovannoli, A., Bernard, R., Boccaccio, I., Boyer, A., et al., 2004. Lamin A and ZMPSTE24 (FACE-1) defects cause nuclear disorganization and identify restrictive dermatopathy as a lethal neonatal laminopathy. *Hum. Mol. Genet.* 13, 2493–2503.

Parsell, D.A., Sauer, R.T., 1989. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J. Biol. Chem.* 264, 7590–7595.

Qin, Z., Buehler, M.J., 2010. Molecular dynamics simulation of the alpha-helix to beta-sheet transition in coiled protein filaments: evidence for a critical filament length scale. *Phys. Rev. Lett.* 104, 198304.

Qin, Z., Kreplak, L., Buehler, M.J., 2009. Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. *PLOS One* 4, e7294.

Rauscher, S., Pomes, R., 2010. Molecular simulations of protein disorder. *Biochem. Cell. Biol.* 88, 269–290.

Rief, M., Fernandez, J.M., Gaub, H.E., 1998. Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* 81, 4764–4767.

Schirmer, E.C., Foisner, R., 2007. Proteins that associate with lamins: many faces, many functions. *Exp. Cell Res.* 313, 2167–2179.

Sheffield, P., Garrard, S., Derewenda, Z., 1999. Overcoming expression and purification problems of RhoGDI using a family of “parallel” expression vectors. *Protein Expr. Purif.* 15, 34–39.

Shumaker, D.K., Lopez-Soler, R.I., Adam, S.A., Herrmann, H., Moir, R.D., et al., 2005. Functions and dysfunctions of the nuclear lamin Ig-fold domain in nuclear assembly, growth, and Emery–Dreifuss muscular dystrophy. *Proc. Natl. Acad. Sci. USA* 102, 15494–15499.

Snow, C.D., Sorin, E.J., Rhee, Y.M., Pandey, V.S., 2005. How well can simulation predict protein folding kinetics and thermodynamics? *Ann. Rev. Biophys. Biomol. Struct.* 34, 43–69.

Sofroniou, M., Schulten, K., 2007. Single-molecule experiments in vitro and in silico. *Science* 316, 1144–1148.

Stierle, V., Couprie, J., Ostlund, C., Krimm, I., Zinn-Justin, S., et al., 2003. The carboxyl-terminal region common to lamins A and C contains a DNA binding domain. *Biochemistry* 42, 4819–4828.

Sugita, Y., Okamoto, Y., 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314, 141–151.

Taimen, P., Pfeleghaer, K., Shimai, T., Moller, D., Ben-Harush, K., et al., 2009. A progeria mutation reveals functions for lamin A in nuclear assembly, architecture, and chromosome organization. *Proc. Natl. Acad. Sci. USA* 106, 20788–20793.

Verstraeten, V.L., Ji, J.Y., Cummings, K.S., Lee, R.T., Lammerding, J., 2008. Increased mechanosensitivity and nuclear stiffness in Hutchinson–Gillford progeria cells: effects of farnesylation inhibitors. *Aging Cell* 7, 383–393.

Wang, N., Tytell, J.D., Ingber, D.E., 2009. Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. *Nat. Rev. Mol. Cell Bio.* 10, 75–82.

Whitmore, L., Wallace, B.A., 2004. Analysis of peptibol sequence composition, implications for vivo synthesis and channel formation. *Eur. Biophys. J.* 33, 233–237.

Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89, 392–400.

Wilson, K.L., Berk, J.M., 2010. The nuclear envelope at a glance. *J. Cell Sci.* 123, 1973–1978.

Wormann, H.J., Bonne, G., 2007. “Laminopathies”: a wide spectrum of human diseases. *Exp. Cell Res.* 313, 2121–2133.

Wormann, H.J., Ostlund, C., Wang, Y., 2010. Diseases of the nuclear envelope. *Cold Spring Harb Perspect. Biol.* 2, a000760.

Yang, S.H., Chang, S.Y., Ren, S., Wang, Y., Andres, D.A., Spielmann, H.P., Fong, L.G., Young, S.G., 2011. Absence of progeria-like disease phenotypes in knock-in mice expressing a non-farnesylated version of progerin. *Hum. Mol. Gen.* 20, 436–444.

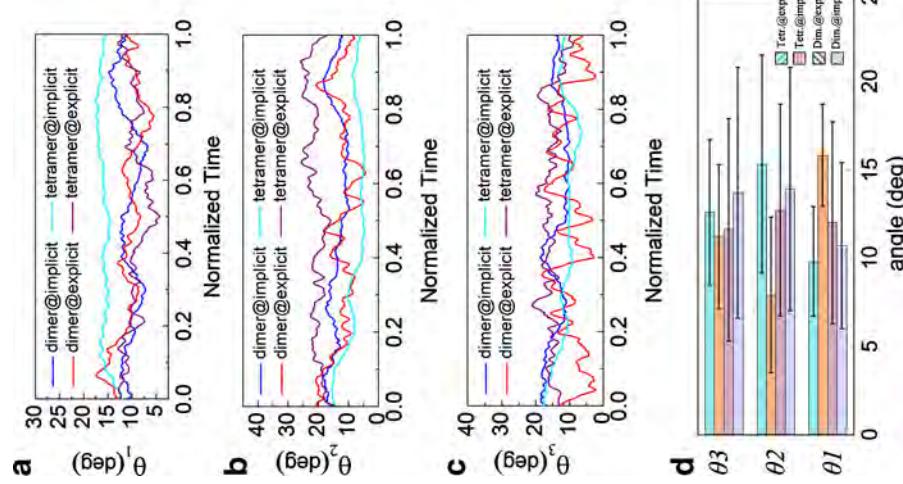
Young, S.G., Fong, L.G., Michaelis, S., 2005. Prelamin A, Zmpste24, misshapen cell nuclei, and progeria—new evidence suggesting that protein farnesylation could be important for disease pathogenesis. *J. Lipid Res.* 46, 2531–2558.

Young, S.G., Meta, M., Yang, S.H., Fong, L.G., 2006. Prelamin A farnesylation and progeroid syndromes. *J. Biol. Chem.* 281, 39741–39745.

Zastrow, M.S., Vlcek, S., Wilson, K.L., 2004. Proteins that bind A-type lamins: integrating isolated clues. *J. Cell Sci.* 117, 979–987.

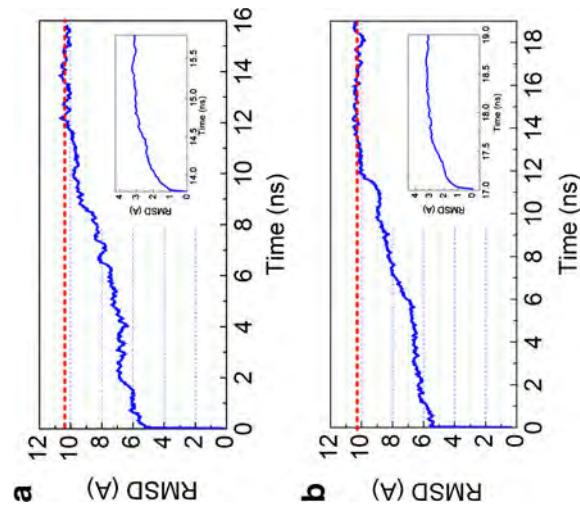
Zhang, S., Iwata, K., Lachmann, M.J., Peng, J.W., Li, S., et al., 2000. The Alzheimer's peptide beta adopts a collapsed coil structure in water. *J. Struct. Biol.* 130, 130–141.

Zhou, R.H., Berne, B.J., Germain, R., 2001. The free energy landscape for beta hairpin folding in explicit water. *Proc. Natl. Acad. Sci. USA* 98, 14931–14936.

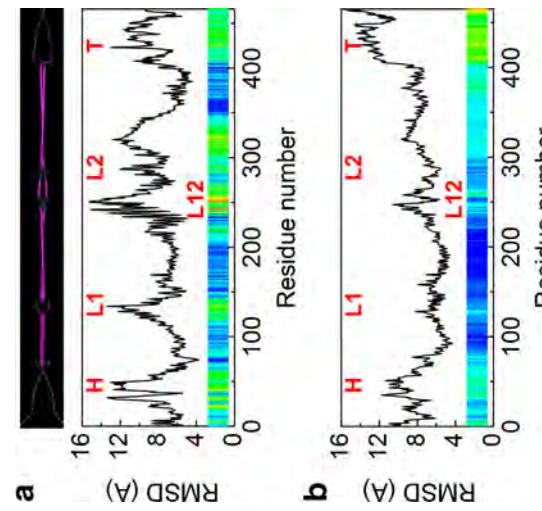


**Fig. 9** Analysis of hinge angles of a dimer and tetramer. Panels **a**, **b** and **c** show how the angles  $\theta_1$ ,  $\theta_2$  and  $\theta_3$  change during the simulation (for geometry see Fig. 5a). The simulation time is normalized by the total equilibration time length (10 ns for the implicit solvent for both dimer and tetramer, 16 ns for the explicit solvent for the dimer, and 19 ns for explicit solvent simulations for the tetramer). Panel **d** summarizes the mean value and standard deviations of the three angles during the simulation. The mean value is given by the main bar, and standard deviations are plotted via error bars

interesting to note that the right part (amino acid numbers 291 to 351) of the 2B segment are more stable than the left part (amino acid numbers 351 to 405), as shown by the peaks in these regions. Experiments have revealed that a stutter defect is located around amino acid number 351. This information provides evidence that the 2B segment is marginally stable from the stutter region on to the beginning of the tail domain, but more flexible before the stutter region. The RMSD value of the tetramer (Fig. 11b) shows that the interaction between two dimers significantly affects the flexibility and stability. As the two *anti*-parallel dimers combine, the overlapped part becomes more stable than the other domains. Moreover, we do not find a peak at the L2 domain, which confirms that this linker interacts with the rod like segment of the other dimer and its flexibility is significantly reduced, reflecting an enhanced stability in this assembly stage.



**Fig. 10** Root mean square displacement (RMSD) analysis in explicit solvent, for the dimer (panel **a**) and the tetramer (panel **b**). The deviation is calculated with respect to the starting configuration, which is the result of equilibration in implicit solvent. Both systems are seen to converge to a constant RMSD value at approximately 12 ns time-scale, with an overall structural deviation of approximately 10 Å. The inserts in each panel depict the RMSD analysis during the last 2 ns of equilibration for the dimer and tetramer, respectively



**Fig. 11** Root mean square displacement (RMSD) analysis during the equilibrium process in explicit solvent (during the period as shown in Fig. 6), for the dimer (panel **a**) and the tetramer (panel **b**). The residues in the range from 1 to 466 refer to the number in one of the polypeptide chains. It corresponds to the average of RMSD values for two chain (dimer) and to four chains (tetramer), respectively. The color bars show the magnitude of RMSD value from low (blue) to high (red). The location of head (H), tail (T) and linker domains (L1, L12, L2) are indicated in the plot. A schematic of the overall dimer structure is shown above panel **a**

## Discussion and conclusions

Based on molecular simulations, we analyzed and compared the structural and dynamical properties of the vimentin dimer and tetramer in implicit and explicit solvent model. In implicit solvent, a long time equilibration was performed in earlier studies reach a minimum energy state. A continued equilibration in explicit solvent as reported in this paper for up to 20 ns reveals that the structural model developed earlier [9] appears to be stable also in explicit solvent, as confirmed in the analyses shown in Figs. 3 and 4, as well as through the RMSD analyses reported in Fig. 10. The analysis of Ramachandran maps as reported in Fig. 5 provides important insight into structural parameters in implicit *vs.* explicit solvents, and could perhaps be used as a method to compare against similar experimental data. The detailed analysis of the number of H-bonds within each dimer and the tetramer as a whole, as well as between the two dimers in the tetramer comparing implicit *vs.* explicit solvent models (see Fig. 6), reveals some differences between the two models, suggesting that the explicit solvent model tends to feature a greater H-bond density. The detailed analysis of the solvent accessible surface area as shown in Fig. 7 suggests that the hidden area between two dimers in explicit solvent is smaller than that of the implicit solvent. The long time-scale simulations provided us with ensembles of conformations, which enabled us to analyze the dynamical properties of the dimer and tetramer around the equilibrium state. This analysis revealed a heterogeneous distribution of the bending stiffness along the molecular axis, where highly flexible hinge-like regions defined by non-alpha-helical linkers connect stiffer regions of the protein, as shown in Figs. 8 and 9.

Overall, our simulations further provided detailed insight into the dynamical properties of the vimentin dimer and tetramer intermediate filament building blocks, which may guide the development of novel coarse-grained models of intermediate filaments that could help in understanding assembly mechanisms. The analysis method used in this paper could be used to gain insight into the structural and dynamical properties of other large biological molecules. However, it should be noted that the analysis reported here has limitations, in particular with respect to accessible time-scales. Accurate calculations can only be performed starting from an equilibrated or near-equilibrated structure, and the simulation time must be sufficient to include enough possible conformations that represent a reasonable sampling of the configurational space. The use of explicit solvents in particular is computational very expensive. In light of these limitations, the implicit solvent calculations show advantages in terms of efficiency at sufficient levels of structural and energetic accuracy.

Future studies of tensile stretching with explicit models could be carried out, albeit these simulations would likely contain millions of atoms due to very large water boxes needed to accommodate the large deformation of the molecules. Other future work could be focused on setting up a coarse-grained *n*-body structural model by treating each dimer as rod-like coiled-coils connected by hinge-like linkers (similar to the suggestion put forth in Fig. 8a), or perhaps using methods that provide enhanced sampling of structural configurations (*e.g.*, replica exchange method) to accelerate the overall simulation speed. The computational expense of full atomistic simulations in explicit solvent is significant in particular if long time-scales are considered. Additional studies could be carried out with advanced time-scale sampling techniques such as replica exchange simulations.

**Acknowledgments** ZQ and MJB acknowledge support by Air Force Office of Scientific Research (AFOSR) (FA9550-08-1-0321) and National Science Foundation (NSF) (MRSEC DMR-0811976). This research was supported by an allocation of advanced computing resources supported by the National Science Foundation (TeraGrid, grant # TG-MSS08030). The authors acknowledge support from the TeraGrid Advanced Support Program. The authors declare no conflict of interest of any sort.

## References

- Alberts B, Johnson A, Lewis J, Raff M, Roberts K et al. (2002) Molecular biology of the cell. Taylor & Francis, New York
- Hermann H, Bar H, Kreplak L, Strelkov SV, Aebi U (2007) Intermediate filaments: from cell architecture to nanomechanics. *Nat Rev Mol Cell Biol* 8(7):562–573
- Wang N, Butler JP, Ingber DE (1993) Mechanotransduction across the cell surface and through the cytoskeleton. *Science* 260 (5111):1124–1127
- Wang N, Stamenovic D (2002) Mechanics of vimentin intermediate filaments. *J Muscle Res Cell Motil* 23(5–6):535–540
- Fudge D, Russell D, Beriault D, Moore W, Lane EB et al. (2008) The intermediate filament network in cultured human keratinocytes is remarkably extensible and resilient. *PLoS ONE* 3(6):e2327
- Qin Z, Buehler MJ, Kreplak L (2010) A multi-scale approach to understand the mechanobiology of intermediate filaments. *J Biomech* 43(1):15–22
- Lewis MK, Nahimy PC, Chen V, Adhikari BB, Wright J et al. (2003) Concentric intermediate filament lattice links to specialized Z-band junctional complexes in sonic muscle fibers of the type I male midshipman fish. *J Struct Biol* 143(1):56–71
- Kreplak L, Hermann H, Aebi U (2008) Tensile properties of single desmin intermediate filaments. *Biophys J* 94(7):2790–2799
- Qin Z, Kreplak L, Buehler MJ (2009) Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. *PLoS ONE* 4(10):e7294
- Dahl KN, Kahn SM, Wilson KL, Discher DE (2004) The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. *J Cell Sci* 117 (20):479–478

11. Wilson KL, Zastrow MS, Lee KK (2001) Lamins and disease: Insights into nuclear infrastructure. *Cell* 104(5):647–650
12. Aebi U, Cohn J, Buhle L, Gerace L (1986) The Nuclear Lamina is a Meshwork of Intermediate-Type Filaments. *Nature* 323 (6088):560–564
13. Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE et al. (2001) Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. *Nat Genet* 27(1):117–120
14. Brown CA, Lanning RW, McKinney KQ, Salvino AR, Chemiske E et al. (2001) Novel and recurrent mutations in lamin A/C in patients with Emery-Dreifuss muscular dystrophy. *Am J Med Genet* 102(4):359–367
15. Bonne G, Mercuri E, Muchir A, Uritzberga A, Becane HM et al. (2000) Clinical and molecular genetic spectrum of autosomal dominant Emery-Dreifuss muscular dystrophy due to mutations of the lamin A/C gene. *Ann Neurol* 48(2):170–180
16. Broers JLV, Hutchison CJ, Ramakers FCS (2004) Laminopathies. *J Pathol* 204(4):478–488
17. Omary MB, Coulombe PA, McLean WH (2004) Intermediate filament proteins and their associated diseases. *N Engl J Med* 351 (20):2087–2100
18. Parry DAD, Strelkov SV, Burkhardt P, Aebi U, Hermann H (2007) Towards a molecular description of intermediate filament structure and assembly. *Exp Cell Res* 313(10):2204–2216
19. Sokolova AV, Kreplak L, Wedig T, Mucke N, Svergun DI et al. (2006) Monitoring intermediate filament assembly by small-angle x-ray scattering reveals the molecular architecture of assembly intermediates. *Proc Natl Acad Sci USA* 103(44):16206–16211
20. Strelkov SV, Hermann H, Lusig N, Geisler N, Ivaninskii S et al. (2001) Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. *J Mol Biol* 306(4):773–781
21. Parry DAD (2006) Hendecad repeat in segment 2A and linker L2 of intermediate filament chains implies the possibility of a right-handed coiled-coil structure. *J Struct Biol* 155(2):370–374
22. Strelkov SV, Hermann H, Geisler N, Wedig T, Zimbelmann R et al. (2002) Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. *EMBO J* 21(6):1255–1266
23. Rafik ME, Doucet J, Briki F (2004) The intermediate filament architecture as determined by X-ray diffraction modeling of hard alpha-keratin. *Biophys J* 86(6):3893–3904
24. Luca S, Yau WM, Leppman R, Tycko R (2007) Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *Biochemistry* 46(47):13505–13522
25. Goldie KN, Wedig T, Mitra AK, Aebi U, Hermann H et al. (2007) Dissecting the 3-D structure of vimentin intermediate filaments by cryo-electron tomography. *J Struct Biol* 158(3):378–385
26. Huang A, Stultz CM (2007) Conformational sampling with implicit solvent models: Application to the PHF6 peptide in tau protein. *Biophys J* 92(1):34–45
27. Zhou RH, Berne BJ (2002) Can a continuum solvent model reproduce the free energy landscape of a beta-hairpin folding in water? *Proc Natl Acad Sci USA* 99(20):12777–12782
28. Nymeyer H, Garcia AE (2003) Simulation of the folding equilibrium of alpha-helical peptides: a comparison of the generalized Born approximation with explicit solvent. *Proc Natl Acad Sci USA* 100(24):13934–13939
29. Lazaridis T, Karplus M (1999) Effective energy function for proteins in solution. *Protein Struct Funct Genet* 35(2):133–152
30. Lazaridis T, Karplus M (1997) “New view” of protein folding reconciled with the old through multiple unfolding simulations. *Science* 278(5345):1928–1931
31. Best RB, Merchant KA, Gopich IV, Schuler B, Bax A et al. (2007) Effect of flexibility and cis residues in single-molecule FRET studies of polyproline. *Proc Natl Acad Sci USA* 104 (48):18964–18969
32. Paci E, Karplus M (2000) Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc Natl Acad Sci USA* 97(12):6521–6526
33. Paci E, Karplus M (1999) Forced unfolding of fibronectin type 3 modules: An analysis by biased molecular dynamics simulations. *J Mol Biol* 288(3):441–459
34. Mackrell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102 (18):3586–3616
35. Nelson MT, Humphrey W, Gursay A, Dalke A, Kale LV et al. (1996) NAMD: A parallel, object oriented molecular dynamics program. *Int J Supercomput Appl* 10(4):251–268
36. Mucke N, Wedig T, Burer A, Marekov LN, Steinert PM et al. (2004) Molecular and biophysical characterization of assembly-starter units of human vimentin. *J Mol Biol* 340(1):97–114
37. Keleti S, Buehler MJ (2008) Geometric confinement governs the rupture strength of H-bond assemblies at a critical length scale. *Nano Lett* 8(2):743–748
38. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99
39. Guzman C, Jeney S, Kreplak L, Kasas S, Kulik AJ et al. (2006) Exploring the mechanical properties of single vimentin intermediate filaments by atomic force microscopy. *J Mol Biol* 360 (3):623–630

# A multi-timescale strength model of alpha-helical protein domains

Theodor Ackbarow<sup>1,2</sup>, Siman Keten<sup>1</sup> and Markus J Buehler<sup>1,3</sup>

<sup>1</sup> Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue Room 1-235A&B, Cambridge, MA, USA

<sup>2</sup> Max-Planck Institute of Colloids and Interfaces, Science Park Golm, 14424 Potsdam, Germany

E-mail: [mbuehler@mit.edu](mailto:mbuehler@mit.edu)

Received 27 August 2008, in final form 12 October 2008

Published 11 December 2008

Online at [stacks.iop.org/JPhysCM/21/035111](http://stacks.iop.org/JPhysCM/21/035111)

## Abstract

Here we report a constitutive model that characterizes the strength of an alpha-helical protein domain subjected to tensile deformation, covering more than ten orders of magnitude in timescales. The model elucidates multiple physical mechanisms of failure in dependence on the deformation rate, quantitatively linking atomistic simulation results with experimental strength measurements of alpha-helical protein domains. The model provides a description of the strength of alpha-helices based on fundamental physical parameters such as the H-bond energy and the polypeptide's persistence length, showing that strength is controlled by energetic, nonequilibrium processes at high rates and by thermodynamical, equilibrium processes at low rates. Our model provides a novel perspective on the strength of protein domains at ultra-slow pulling speeds relevant under physiologic and experimental conditions.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

Alpha-helical protein domains, together with beta-sheets and tropocollagen molecules, represent one of the most abundant protein structures found in biology. In addition to being part of larger protein structures (such as in enzymes and other globular proteins), alpha-helical protein domains also play an important mechanical and structural role in biology. For example, alpha-helix networks in intermediate filaments have been shown to provide mechanical integrity to cells, and also to support biological processes that involve mechanical signalling such as mechanotransduction or mechanosensation to regulate gene activation [1–4]. Therefore, in order to advance our understanding of such biological processes, a quantitative understanding of the mechanical behavior of alpha-helices is crucial. In addition to medical and biological applications, a detailed understanding of alpha-helices and alpha-helix based protein networks and their resulting mechanical properties could possibly lead to the creation of *de novo* synthetic alpha-helix based materials [5, 6].

<sup>3</sup> Author to whom any correspondence should be addressed.

The mechanical properties of alpha-helices must be understood through disparate timescales, reaching from picoseconds (e.g. during injuries, trauma, mechanical shock) to seconds and more (e.g. during regular physiological cellular processes) [3, 4, 7, 8]. However, currently there exists no model that describes the mechanical strength behavior of alpha-helical protein domains that considers associated physical mechanisms through this exhaustive range of timescales. Experiments have been carried out at relatively slow pulling rates (of the order of nanometers per second), and computer simulations (e.g. molecular dynamics simulations) have been carried out at much faster deformation rates (of the order of meters per second). The results of such experimental and computational studies have not yet been integrated. Understanding the behavior of proteins over multiple timescales and deriving the behavior at one timescale from the known behavior at another timescale is crucial to enable future biological research and to develop engineering design methods to create *de novo* biological protein materials. Currently no model has been reported that can predict experimentally accessible pulling speeds based on the analysis of molecular dynamics (MD) simulation results.

Moreover, MD simulation studies typically cannot be directly extrapolated towards lower pulling speeds, since they predict unphysical phenomena such as negative strength values.

Here we resolve this issue by providing a self-consistent approach that allows us to predict the strength of alpha-helices over more than ten orders of magnitude in timescales, quantitatively linking atomistic simulation results with experimental results, based on fundamental physical parameters that include the energy and geometry of H-bonds (HBs) and the persistence length of the protein's backbone. The model captures the behavior of alpha-helices from 'slow' natural biological processes up to mechanical shock as it appears in accidents and injuries.

## 2. Computational approach

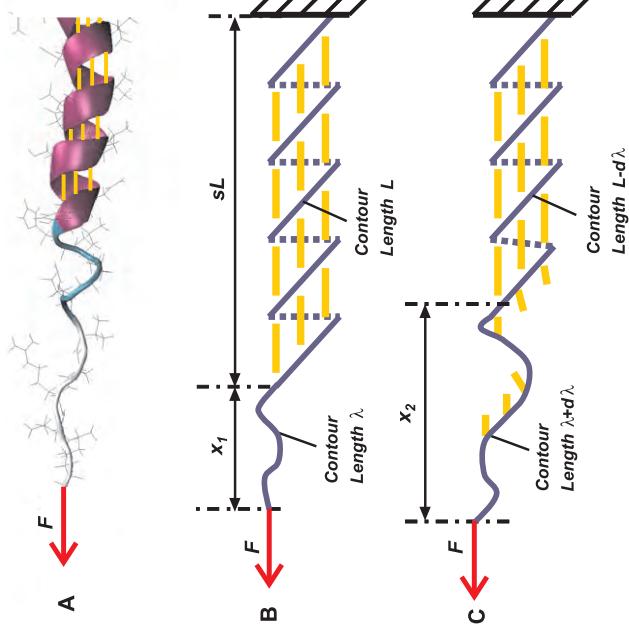
To demonstrate the link between computer simulation and experiment, we utilize a set of MD simulations that were reported in earlier simulation studies [9]. For convenience, here we review details of the computational atomistic modeling approach. For all atomistic simulations, we use a classical MD approach, implemented in the MD program NAMD [10] using the CHARMM22 force field [11]. All simulations are performed at a temperature of 300 K (*NVT* ensemble, temperature control using a Berendsen thermostat), with a time step of 1 fs. Careful energy minimization and finite temperature equilibration of all structures are simulated before the protein domain is mechanically loaded. The protein structure obtained from the Protein Data Bank (PDB identifier 1gk6, part of human vimentin intermediate filament) is solved completely in a TIP3 water skin. In all cases studied here, the entire protein is embedded in water, before and during the deformation of the protein. This is essential to capture the correct HB rupture dynamics.

To apply forces to the molecule in order to induce deformation, we use steered molecular dynamics (SMD) [12], with SMD spring constant  $k_{\text{SMD}} = 10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ . We obtain force versus displacement data by monitoring the time-averaged applied force ( $F$ ) and the position of the atom that is pulled ( $x$ ) over the simulation time.

To apply load,  $\text{C}_\alpha$  atoms at one end are fixed and the force is applied on the  $\text{C}_\alpha$  atom at the other end in the AH structure, with a pulling speed  $v$ . The tensile boundary conditions chosen for the AH domain are closest to the physiological conditions. Several other boundary conditions have been used (changing fixed and pulled atoms, pulling at different convolutions). No changes in the rupture forces have been observed, suggesting that the results reported here are robust with respect to changes in the boundary conditions.

## 3. Theoretical modeling and results

A cartoon of the AH protein and a schematic diagram of the tensile load boundary conditions used to study the rupture mechanism are shown in figure 1(A). As reported in previous work, MD simulations of AHs in explicit solvent were performed over four orders of magnitude of pulling speeds (from 0.05 to 100 m s<sup>-1</sup> [9]). The rupture force of the AH



**Figure 1.** (A) The atomistic-scale protein structure of a single alpha-helix (AH) from a vimentin coiled-coil dimer. The helical backbone is stabilized by parallel arrangements of hydrogen bonds (HBs, yellow dashed lines). (B), (C) A schematic model system of an AH strained by an external force before and after onset of rupture, showing the process of releasing a segment of backbone polypeptide due to the rupture of HBs, thereby increasing the contour length of the free end entropic chain by  $d\lambda$ .

structure, identified at the point of breaking of the first HBs, is plotted as a function of the protein domain's lifetime  $\tau$  in figure 2(A). When the system is not in equilibrium, as is the case for high deformation rates, the relation between  $\tau$  and the applied force  $f$  can be described by a simple Bell model [13]:

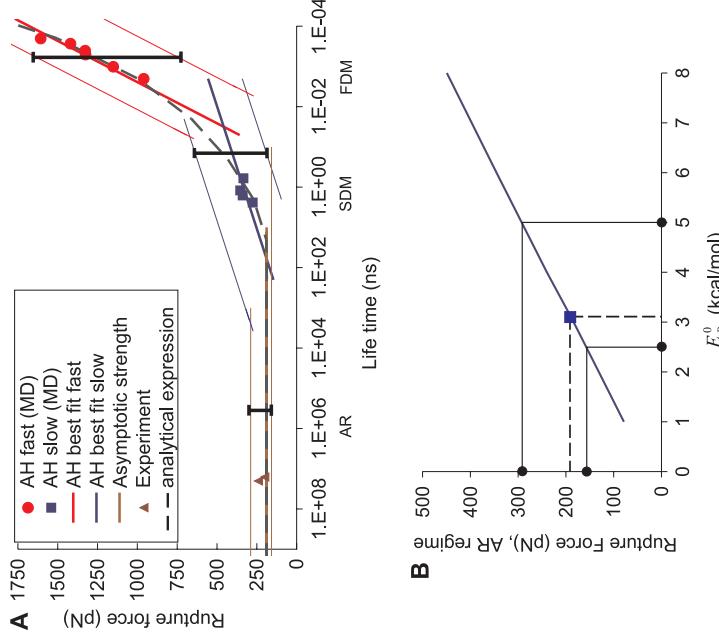
$$\tau = \omega_0^{-1} \exp \left( \frac{E_b - f x_b \cos(\theta)}{k_B T} \right), \quad (1)$$

where  $E_b$  is the energy barrier of HB breaking, and  $x_b$  is the distance between the equilibrium state and the transition (=rupture) state of the protein domain (note that  $v = \Delta x / \Delta t = x_b / \tau$ , where  $v$  is the externally applied pulling speed). Further, the parameter  $\theta \approx 16^\circ$  describes the angle between the applied force  $f$  and the orientation of the HBs,  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $\omega_0 = 10^{13} \text{ s}^{-1}$  is the natural frequency of bond vibration.

It is noted that, in addition to the phenomenological model used here, other stochastic models exist that link timescales and pulling speeds to bond breaking forces; for a description of other models we refer to the literature [14–22]. The force as a function of timescale  $\tau$  and the energy landscape parameters (ELPs,  $E_b$  and  $x_b$ ) is given by

$$f(\tau, E_b, x_b, \theta) = (x_b \cos(\theta))^{-1} [E_b - k_B T \ln(\omega_0 \tau)]. \quad (2)$$

(For detailed explanations of these equations see [9].) For a given pair of ELPs, equation (2) leads to a straight line in the  $f$ – $\ln(\tau)$  space. Direct MD simulation studies in explicit water confirm this predicted behavior, however, we observe



**Figure 2.** (A) Rupture force versus lifetime of the AH system at the onset of failure (=strength properties), including all three regimes over more than ten orders of magnitude of timescales. MD simulation results (as reported in [9]) suggest a change in mechanism from the fast deformation mode (FDM) to the slow deformation mode (SDM) on increasing the timescales. At approximately 350 pN the effective energy barriers under the applied force in the Bell model are comparable, and therefore mark the transition between FDM and SDM mechanisms. At longer timescales there is another change in deformation mechanism from the SDM to the asymptotic regime (AR), predicted here at a timescale of approximately 100 ns when  $f_{\text{AR}} > f_{\text{FDM}}$ . Experimental results confirm this prediction. Thin lines show the strength behavior for a broad range of HB energy values from 2.5 to 5 kcal mol<sup>-1</sup> (marking error bars for uncertainties in H-bond energy). The dashed line represents a closed analytical expression from the power law fit over all three regimes, provided in equation (1.1). (B) Dependence of the critical rupture force on  $E_B^0$ , in the AR. The strength of the system near equilibrium conditions (AR) depends linearly on  $E_B^0$  (this parameter determines the energy release rate  $\gamma_s$ ). The specific value of  $E_B^0$ , usually found in a range between 1 and 8 kcal mol<sup>-1</sup>, varies between different solvent conditions and the specific sequence of the protein domain.

(short timescales, FDM) is due to the fact that HB breaking in the protein remains localized. This is because pulling occurs faster than the ability of the protein to mediate HB breaking induced ‘plastic’ deformation. In the SDM regime, however, pulling is slow enough that entire convolutions rupture under the applied force, leading to effectively higher energy barriers for unfolding [9].

At increasing timescales in the SDM the Bell model prediction leads to negative forces, an unphysical prediction. Furthermore, experimental values [23, 24] clearly do not lie on an extension of the slope predicted from the SDM regime, and rather suggest that the  $f - \ln(v)$  curve approaches an asymptotic zero slope (see figure 2(A)). Could the Bell model be used to explain this behavior at vanishing pulling rates? Adopting the Bell model to describe this behavior would lead to an increase of  $x_B$  (since  $x_B$  controls the slope of the  $f - \ln(\tau)$  curve), approaching infinity for slopes approaching zero. It is noted that in other models (e.g. the microscopic theory [19, 25]) a similar approach has been taken, where the value of  $x_B$  is defined as a function of pulling speed (equivalently, the timescale), leading to a continuous change of the slope of the  $f - \ln(v)$  curve.

The approach of  $x_B$  to extremely large values is, however, unphysical, since the transition point  $x_B$  can not be larger than the finite contour length of the protein domain. This suggests that another mechanism must determine the protein rupture force. The key to understand this change in mechanism is the realization that at sufficiently long timescales the deformation of the system goes through equilibrium and is no longer controlled by a statistically activated process as described in the Bell model or equations (1)–(2). Thus the strength does not depend on the timescale of loading beyond a critical  $\tau_{\text{crit}}$ , and is independent of pulling rate for very long timescales.

At long timescales  $\tau > \tau_{\text{crit}}$  entropic effects that stem from conformational changes of the polypeptide chains are activated and the strength is characterized by a free energy release rate condition. A similar approach has recently been reported for beta-sheets in [26]. Here we develop a model specific to AH protein domains. The aim here is to find the critical force, that will initiate rupture of HBs in an alpha-helix at quasi-equilibrium deformation rates. Similar to the Griffith condition used to predict the onset of fracture in crystals [27], the free energy released by freeing polypeptide chains from their geometric confinement in helical convolutions must equal the energy required to break these HBs. The free energy balance condition at the onset of fracture requires that  $G = -(A_2 - A_1 - F\delta)/d\lambda = \gamma_s$ , where  $\gamma_s$  denotes energy released by rupture of HBs per unit crack advance,  $F\delta$  is the work done by the external force on the system, and  $A_1$  and  $A_2$  are the initial and final free energies of the protein backbone as determined from the worm-like chain elasticity theory. The free energy of the system before and after rupture is given as

$$A_1 = \lambda A_{\text{WLC}} - \gamma_s L + L A_{\text{FOLD}} \quad (3)$$

$$A_2 = (\lambda + d\lambda) A_{\text{WLC}} - \gamma_s (L - d\lambda) + (L - d\lambda) A_{\text{FOLD}}, \quad (4)$$

**Table 1.** Comparison and summary of the three deformation regimes (FDM = fast deformation mode, SDM = slow deformation mode, AR = asymptotic regime) with their characteristic physical parameters and numerical values.

Mechanism, associated pulling speeds (in m s <sup>-1</sup> )	Timescale (in ns)	Force levels (in pN)	Physical parameters	Controlling physical mechanism and explanation
AR $v < 0.001$	$\tau > 100$	$F < 200$	$\gamma_s = 0.91 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ $\xi_p = 4 \text{ \AA}$	Thermodynamical free energy release rate through equilibrium (asymptotic strength model)
SDM $0.001 < v < 0.4$	$0.05 < \tau < 100$	$200 < F < 350$	$E_b^{\text{SDM}} = 11.1 \text{ kcal mol}^{-1}$ $x_b^{\text{SDM}} = 1.2 \text{ \AA}$	Simultaneous rupture of HBs in one convolution, activated statistical process (Bell)
FDM $v > 0.4$	$\tau < 0.05$	$F > 350$	$E_b^{\text{FDM}} = 4.87 \text{ kcal mol}^{-1}$ $x_b^{\text{FDM}} = 0.2 \text{ \AA}$	Sequential rupture of HBs, activated statistical process (Bell)

where  $A_{\text{WLC}} = \int_0^\alpha F_{\text{WLC}}(\alpha) d\alpha$  is the free energy state (energy per length) of the already unfolded free segments of the protein and  $A_{\text{FOLD}} = \int_0^s F_{\text{WLC}}(\alpha) d\alpha$  is the free energy state of the folded segment of the chain. Hereby  $\alpha$  is equal to the ratio of the end-to-end length of the free chain to its contour length  $\alpha = x/\lambda$ , equivalent to mechanical stretch, and the parameter  $s$  denotes the ratio of the end-to-end length of the alpha-helix to its contour length,  $L$  (the physical meaning of this parameter is that it describes how much contour length is stored per unit length of alpha-helix). We refer the reader to figure 1 for an illustration of the definition of variables. Equation (4) illustrates the interplay between entropic energy release in the stretched and relaxed segments of the chain with energetics of HB rupture, thereby coupling two key physical aspects of the protein unfolding problem. The energy contribution from the external force is given as

$$\delta W_F = -F(\alpha - s)d\lambda. \quad (5)$$

Hence the critical condition for HB rupture can be given as

$$A_{\text{WLC}}(\alpha_{\text{cr}}) + F(s - \alpha_{\text{cr}}) + \gamma_s - A_{\text{FOLD}}(s) = 0, \quad (6)$$

where  $\alpha_{\text{cr}}$  is the critical stretch level that initiates rupture. The strength regime described by equation (6) is referred to as the asymptotic regime (AR), and the force prediction is then found by the WLC model, through  $f_{\text{AR}} = F_{\text{WLC}}(\alpha_{\text{cr}})$ , leading to

$$f_{\text{AR}} = \frac{k_B T}{4\xi_p} [(1 - \alpha_{\text{cr}})^{-2} + 4\alpha_{\text{cr}} - 1]. \quad (7)$$

With the core theory established, we can now substitute quantitative values for the parameters. The parameter  $\gamma_s$  describes the HB energy stored per unit length of AH and can be obtained from

$$\gamma_s = \frac{E_b^0}{L_0}, \quad (8)$$

where  $E_b^0$  is the dissociation energy of a single bond and  $L_0 = 0.33 \text{ nm}$  is the distance between adjacent HBs along the length of the helix. The parameter  $s = 0.45$  can be estimated from atomistic simulations of the deformation mechanics of alpha-helices, where the unfolded length of the molecule can easily be calculated to find the ratio with initial end-to-end distance.

These values are also in excellent agreement with the well-established alpha-helix pitch of  $5.4 \text{ \AA}$  per convolution [28, 29]. We further note that the fracture model is independent of the size of the macromolecule and the helical domain. This is because the initial unfolded contour length does not influence the strength prediction.

Combining all three mechanisms (FDM, SDM, AR), the strength of an AH domain is

$$F(\tau; x_b^{\text{FDM}}, E_b^{\text{FDM}}, x_b^{\text{SDM}}, E_b^{\text{SDM}}, \theta, \xi_p, \gamma_s) = \max \left\{ \begin{array}{l} f_{\text{FDM}}(\tau; x_b^{\text{FDM}}, E_b^{\text{FDM}}, \theta) \\ f_{\text{SDM}}(\tau; x_b^{\text{SDM}}, E_b^{\text{SDM}}, \theta) \\ f_{\text{AR}}(\xi_p, \gamma_s(E_b^0)) \end{array} \right\}. \quad (9)$$

The functions  $f_{\text{FDM}}$  and  $f_{\text{SDM}}$  can be calculated from equation (2), and  $f_{\text{AR}}$  can be calculated from equations (6)–(8). We estimate  $E_b^0$  from the MD simulation results in the SDM, where the 3.6 HBs in one convolution break simultaneously, thus  $E_b^0 = E_b^{\text{SDM}}/3.6 = 3.1 \text{ kcal mol}^{-1}$ , and therefore  $\gamma_s = 0.91 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ . This relation between  $E_b^0$  and  $E_b^{\text{SDM}}$  shows the ability of our model to link directly between the AR and results in the SDM. A similar link can be established between the AR and FDM, where the energy barrier found in FDM typically directly corresponds to the HB energy, since rupture occurs sequentially.

In summary, once the structural parameters (e.g. the ratio  $s$ , the persistence length  $\xi_p$ , the distance of HBs  $L_0$ ) and energetic parameters (e.g. the energy of individual HBs,  $E_b^0$ ) are determined from MD simulations, the force level in the AR can be calculated based on the theoretical link developed here, by solving equations (6)–(8). In analogy, the inverse calculation is possible as well. Rupture forces as they would occur at very high pulling speeds (e.g. as they appear during injuries) can be calculated from data generated at slow pulling speeds, for example through experimental analysis, by applying equation (2). Being able to calculate and predict the behavior at one timescale from observations in another timescale, as achieved in this model, underlines the coupled, multi-timescale character of our theory.

The value of  $E_b^0$  determined from the MD simulation studies is in good agreement with earlier experimental and simulation results [30], where  $E_b^0$  was reported to be in the

range of 3–6 kcal mol<sup>−1</sup>. We choose the persistence length of a polypeptide chain as suggested from both experiment and theory to be  $\xi_P = 4 \text{ \AA}$  [31]. Based solely on these two parameters,  $E_b^0$  and  $\xi_P$ , the force in the AR is calculated to be  $\approx 189 \text{ pN}$ . The AR regime is reached at a critical timescale of 100 ns (or equivalently at pulling speeds  $v < 0.001 \text{ m s}^{-1}$ ), when  $f_{\text{AR}} > f_{\text{SDM}}$ . The strength value of  $f_{\text{AR}}$  is plotted in figure 2(B) as a function of the HB energy  $E_b^0$ .

In order to facilitate the direct application of our model as a constitutive equation in a multi-scale simulation approach (e.g. as a strength model), we have fitted the results to an empirical relation that provides a single (empirical) mathematical expression that interpolates through all timescales (and thus all modes of deformation mechanisms). The force as a function of timescale is expressed as

$$F(\tau) = \xi_1(\tau) + h(\tau)(\xi_0 - \xi_1(\tau)), \quad (10)$$

where  $\xi_0$  describes the force level in the AR regime, given by  $f_{\text{AR}}$ . The use of a smooth Heaviside function  $h(\tau) = 1/(1 + (\tau_c/\tau)^k)$  enables us to describe the transition from AR to the rate dependent regimes (FDM and SDM). In this Heaviside function  $\tau_c$  characterizes the timescale of the transition, and  $k$  determines the sharpness of the transition. The Heaviside function approaches  $h(\tau) = 1$  for  $\tau \ll \tau_c$ , and is zero for  $\tau \gg \tau_c$ . The strength dependence on timescale in the FDM and SDM regimes is approximated using a power law of the form  $\xi_1(\tau) = b_1 \tau^{b_2}$ . The parameters in equation (10) are fitted to reproduce the overall behavior shown in figure 2, leading to  $\tau_c = 7.4 \text{ ns}$ ,  $k = 1$ ,  $\xi_0 = 189 \text{ pN}$ ,  $b_1 = -330 \text{ pN}$  and  $b_2 = -0.18$ . The complete expression is

$$F(\tau) = b_1 \tau^{b_2} + \frac{1}{1 + (\tau_c/\tau)^k} (\xi_0 - b_1 \tau^{b_2}). \quad (11)$$

The fit to the simulation and experimental results is shown in figure 2(A) as the dashed line.

The model developed above (and the numerical interpolation given in equations (10) and (11)) is validated through quantitative comparison with experimental results. Experimental results of stretching and breaking single AH domains [23, 24] (with a length of less than 100 Å) report forces between 140 and 240 pN during unfolding. Figure 2(A) summarizes the described regimes and shows a quantitative comparison between the model prediction and MD simulation results as well as experimental results. In addition to the values used in this study that were based on earlier MD results, an envelope curve for  $E_b^0$  ranging from 2.5 to 5 kcal mol<sup>−1</sup> is included to illustrate how the predictions change under variations of the energy of HBs. We note that other experimental results [31–35] (not shown in figure 2) that consider AH spectrum repeats lie slightly below the predicted force range, of the order of 50 pN, which would require extremely low values of  $E_b^0 \approx 1 \text{ kcal mol}^{-1}$ . A possible explanation for this behavior could be the difference in the observed unfolding mechanism, which is the unfolding of the anti-parallel coiled-coil repeat instead of rupture of individual HBs of an AH domain. For instance, in one of the studies  $x_b$  was estimated to be 15 Å [31–35], which is ten times higher

than the  $x_b$  for a single HB, thus suggesting an alternative rupture mechanism.

#### 4. Summary and discussion

The most important contribution of this paper is the development of a constitutive model (equations (9) and (11)) that describes the strength properties of AH protein domains over more than 10 orders of magnitude of timescales. Up until now such a model has not been reported, and to the best of our knowledge this model is the first to quantitatively link MD simulation results [9] and experimental AH strength values [23, 24] in a simple physical model as shown in figure 2(A). An important feature of the model reported in equation (9) is that it only includes basic parameters of the protein structure, that is, the HB energy and geometry, as well as persistence length. The strength properties of the AH protein domain, a universally found biological protein structure, are controlled by different mechanisms at distinct timescales, with strong strengthening under faster rates (shorter timescales).

According to our model, the strength at very slow pulling rates is controlled by energetics of HB rupture and entropic effects of the unfolding polypeptide backbone, and not by a continuously changing energy barrier that moves along the reaction coordinate  $x_b$  as suggested in the microscopic theory. Changes in the reaction coordinate  $x_b$  are only observed at relatively fast pulling rates, where it can be directly linked to changes in the physical mechanism of rupture (that is, the change from FDM to SDM as reported from MD simulation studies). Further, we have shown here that the deformation mechanisms that appear in MD simulations are likely different from those that appear in experiments. However, even though carried out at much faster pulling speeds, MD simulations allow us to determine basic parameters such as  $E_b^0$ . These parameters can then be used to predict force levels that appear *in vivo* or in experimental studies. This has been achieved here through the introduction of the model that characterizes the strength of alpha-helical protein domains in the AR regime. Our study could motivate new experiments, in particular those that would provide a systematic variation of deformation rates to probe the transitions between the regimes described here.

Since our model is derived from fundamental principles, such as the rupture energetics of HBs and entropic effects (which appear universally in almost any protein structure), it should be applicable to other protein structures (e.g. amyloid beta-helices or tropocollagen molecules). Our model becomes specific to a particular protein structure solely through parameters that define the geometry, such as  $L_0$ ,  $s$ ,  $\theta$  and  $x_b$ , as well as related energetic parameters such as  $E_b^0$  and  $\gamma$ . Environmental conditions such as salt concentration, pH and the exposure to water (e.g. due to geometric confinement into larger protein structures that shield from the direct exposure to water) are captured by the value of  $E_b^0$ , which describes the energy necessary for breaking a single HB. Different protein structures with different geometries in different environments feature different levels of  $E_b^0$ , and as a consequence the strength values as well as the timescale at which changes in deformation mechanism appear are expected to vary. The variation of

the strength, however, remains within relatively small error margins as shown in figure 2(A)/(B) (see error bars that show the change of strength across all regimes due to changes of the HB energy from  $2.5$  to  $5$   $\text{kcal mol}^{-1}$ , and the analysis in figure 2(B) that shows the variation of strength for a broad range of values of the HB energy).

### Acknowledgments

This research was supported by AFOSR (program manager Dr Les Lee and by the Army Research Office, grant number W911NF-06-1-0291 (program officer Dr Bruce LaMattina). TA acknowledges support from the German National Academic Foundation and the Hamburg Foundation for International Research Studies (Germany). We thank Professor Lothar Gaul (University Stuttgart) and Professor Reinhard Lipowsky (Max-Planck Institute of Colloids and Interfaces Potsdam) for their continuous interest and support of our work. The authors state that they have no competing financial interests.

### References

- [1] Alberts B *et al* 2002 *Molecular Biology of the Cell* (London: Taylor and Francis)
- [2] Gruber M and Lupas A N 2003 Historical review: another 50th anniversary—new periodicities in coiled coils *Trends Biochem. Sci.* **28** *679–85*
- [3] Moir R D and Spann T P 2001 The structure and function of nuclear lamins: implications for disease *Cell. Mol. Life Sci.* **58** *1748–57*
- [4] Wilson K L, Zastrow M S and Lee K K 2001 Lamins and disease: insights into nuclear infrastructure *Cell* **104** *647–50*
- [5] Bryson J W *et al* 1995 Protein design—a hierarchical approach *Science* **270** *935–41*
- [6] Kirshenbaum K, Zuckermann R N and Dill K A 1999 Designing polymers that mimic biomolecules *Curr. Opin. Struct. Biol.* **9** *530–5*
- [7] Kim S and Coulombe P A 2007 Intermediate filament scaffolds fulfill mechanical, organizational, and signaling functions in the cytoplasm *Genes Dev.* **21** *1581–97*
- [8] Herrmann H *et al* 2007 Intermediate filaments: from cell architecture to nanomechanics *Nat. Rev. Mol. Cell Biol.* **8** *562–73*
- [9] Ackbarow T *et al* 2007 Hierarchies, multiple energy barriers and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains *Proc. Natl. Acad. Sci. USA* **104** *16410–5*
- [10] Nelson M T *et al* 1996 NAMD: A parallel, object oriented molecular dynamics program *Int. J. Supercomput. Appl. High Perform. Comput.* **10** *251–68*
- [11] MacKerell A D *et al* 1998 All-atom empirical potential for molecular modeling and dynamics studies of proteins *J. Phys. Chem. B* **102** *3586–616*
- [12] Lu H *et al* 1998 Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation *Biophys. J.* **75** *662–71*
- [13] Bell G I 1978 Models for specific adhesion of cells to cells *Science* **200** *618–27*
- [14] Evans E A and Calderwood D A 2007 Forces and bond dynamics in cell adhesion *Science* **316** *1148–53*
- [15] Evans E 2001 Probing the relation between force—lifetime—and chemistry in single molecular bonds *Annu. Rev. Biophys. Biomol. Struct.* **30** *105–28*
- [16] Evans E B 1999 Looking inside molecular bonds at biological interfaces with dynamic force spectroscopy *Biophys. Chem.* **82** *83–97*
- [17] Merkell R *et al* 1999 Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy *Nature* **379** *50–3*
- [18] Evans E and Ritchie K 1997 Dynamic strength of molecular adhesion bonds *Biophys. J.* **72** *1541–55*
- [19] Dudko O K, Hummer G and Szabo A 2006 Intrinsic rates and activation free energies from single-molecule pulling experiments *Phys. Rev. Lett.* **96** *108101*
- [20] Makarov D E 2007 Unraveling individual molecules by mechanical forces: theory meets experiment *Biophys. J.* **92** *4135–6*
- [21] West D K, Olmsted P D and Paci E 2006 Mechanical unfolding revisited through a simple but realistic model *J. Chem. Phys.* **124** *154909*
- [22] Erdmann T and Schwarz U S 2004 Stability of adhesion clusters under constant force *Phys. Rev. Lett.* **92** *108102*
- [23] Lantz M A *et al* 1999 Stretching the alpha-helix: a direct measure of the hydrogen-bond energy of a single-peptide molecule *Chem. Phys. Lett.* **315** *61–8*
- [24] Kageshima M *et al* 2001 Insight into conformational changes of a single alpha-helix peptide molecule through stiffness measurements *Chem. Phys. Lett.* **343** *77–82*
- [25] Dudko O K *et al* 2007 Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins *Biophys. J.* **92** *4188–95*
- [26] Keten S and Buelter M J 2008 Asymptotic strength limit of hydrogen-bond assemblies in proteins at vanishing pulling rates *Phys. Rev. Lett.* **100** *198301*
- [27] Griffith A A 1920 The phenomenon of rupture and flows in solids *Phil. Trans. R. Soc. A* **221** *163–98*
- [28] Brändén C-I and Tooze J 1999 *Introduction to Protein Structure* 2nd edn (New York: Garland) chapter 3
- [29] Voet D and Voet J G 2004 *Biochemistry* (New York: Wiley)
- [30] Sheu S-Y *et al* 2003 Energistics of hydrogen bonds in peptides *Proc. Natl. Acad. Sci.* **100** *12683–7*
- [31] Rief M *et al* 1999 Single molecule force spectroscopy of protein repeats: low unfolding forces in helix bundles *J. Mol. Biol.* **286** *553–61*
- [32] Law R *et al* 2004 Influence of lateral association on forced unfolding of antiparallel spectrin heterodimers *J. Biol. Chem.* **279** *16410–6*
- [33] Lenné P F *et al* 2000 States and transitions during forced unfolding of a single spectrin repeat *FEBS Lett.* **476** *124–8*
- [34] Law R *et al* 2003 Cooperativity in forced unfolding of tandem spectrin repeats *Biophys. J.* **84** *533–44*
- [35] Law R *et al* 2003 Pathway shifts and thermal softening in temperature-coupled forced unfolding of spectrin domains *Biophys. J.* **85** *3286–93*

Intermediate filaments (IFs) are the key components of cytoskeleton in eukaryotic cells and are critical for cell mechanics. The building block of IFs is a coiled-coil alpha-helical dimer, consisting of several domains that include linkers and other structural discontinuities. One of the discontinuities in the dimer's coiled-coil region is the so-called 'stutter' region. The stutter is a region where a variation of the amino acid sequence pattern from other parts of the alpha-helical domains of the protein is found. It was suggested in earlier works that due to this sequence variation, the perfect coiled-coil arrangement ceases to exist. Here, we show using explicit water molecular dynamics and well-tempered metadynamics that for the coil2 domain of vimentin IFs the stutter is more stable in a non-alpha-helical, unfolded state. This causes a local structural disturbance in the alpha helix, which has a global effect on the nanomechanics of the structure. Our analysis suggests that the stutter features an enhanced tendency to unfolding even under the absence of external forces, implying a much greater structural instability than previously assumed. As a result it features a smaller local bending stiffness than other segments and presents a seed for the initiation of molecular bending and unfolding at large deformation.

5

## Introduction

Intermediate filaments (IFs) are a major building material of the cytoskeleton in eukaryotic cells, and are crucial in the determination of cell mechanical properties (Hutchison 2002; Herrmann and Aebi 2004; Herrmann et al. 2007). IFs are crucial in defining key functions of cells such as cell migration, cell division and mechanotransduction, and it is now understood that they play a critical role similar to other cytoskeleton elements such as microtubules and microfilaments. Notably and pointing to their rather important role as mechanical elements, IFs have also been referred to as the 'safety belts of cells', and have been demonstrated to play a role in preventing exceedingly large cell stretch (Wang et al. 1993; Wang and Stamenovic 2002; Fudge et al. 2008). A range of studies published over the past years focused on the mechanical signature of IFs, which have suggested that they can sustain extremely large deformation of up to several 100% (Lewis et al. 2003; Herrmann et al. 2007; Kreplak et al. 2008). It was also found that due to severe stiffening, the tangent modulus of IFs increases manifold during deformation, which clearly could be crucial for providing mechanical resistance against large stretching. In addition to the cell's cytoskeleton, IFs are also found in the cell nucleus in the form of lamin IF, where they form a dense mesh-like network providing mechanical integrity and biochemical<sup>30</sup> 35 40 45 50

functions at the cytoskeleton-chromatin interface (Wilson et al. 2001; Dahl et al. 2004; Ackbarow et al. 2009). Numerous diseases associated with IFs were identified which point to the clear importance of material factors in the initiation and progression of these disorders (Brenner et al. 2001; Omary et al. 2004; Buehler and Yung 2009). The basic building block of all IF proteins is a particularly stable molecular configuration of H-bonded alpha-helix-based protein structures called coiled coils (Alberts 2002; Rose and Meier 2004). This arrangement of two alpha helices into coiled-coil domains provides increased strength to the structures. In alpha-helix-based-coiled coils within a single heptad repeat (abcdefg periodicity), some positions are occupied by hydrophilic (polar) and some are occupied by hydrophobic (non-polar) residues (Alanine (ALA), Isoleucine (ILE), Leucine (LEU)...) (Strelkov et al. 2003; MacCallum et al. 2007; Buehler and Ackbarow 2008). The hydrophobic residues constitute the reason why these proteins assemble into coiled-coil structures as they ensure minimal contact of hydrophobic domains with water.

Vimentin IFs are the most widely distributed types among all IFs (Alberts 2002) and will be the focal point of the study reported here. The structure is made up of 310 amino acid residues consisting of four segments: I1, I2, 1A, 1B, 2A and 2B divided by three linkers: I1-I12-I2 as shown in Figure 1. The I1 and I2 domains with their 1A and 1B domains are shown in Figure 2.

\*Comments and questions can be sent to [ambushlan@mit.edu](mailto:ambushlan@mit.edu).

Figure 1(A). Segment 2A (residues 264–282) is the shortest one among all segments. It assumes the configuration of a nearly parallel alpha-helical bundle, in agreement with earlier suggestions (Parry et al. 2007; Qin et al. 2009; Nicolet et al. 2010). Segment 2B with 115 amino acids (residues 291–405) features alpha-helical coiled-coil geometry for the major segment. One of the discontinuities found in heptad repeats of an alpha-helical coiled-coil protein is the ‘stutter’ region. As a result it has been suggested that the molecular structure may be altered in this region and deviate from the perfectly coiled regions in the rest of the protein filament. The 2B segment of vimentin contains the stutter region as visualised in Figure 1(A) (Brown et al. 1996; Strelkov et al. 2001; Ackbarow and Buehler 2009). It can be seen that the stutter is located at the end of the eighth heptad, in the vicinity of residue 351 (Strelkov 2002). An analysis of the amino acid sequence revealed that the nature of the vicinity of the stutter region is hydrophilic. Earlier work suggested that this results in a parallel arrangement for the two coils in the alpha helix rather than a coiled-coil configuration.

## Materials and methods

### Molecular model construction

Recent advances in identifying the structure of vimentin dimers with atomistic details (Strelkov et al. 2002; Nicolet et al. 2010), combined with protein simulation tools, now

enable us to build molecular mechanics models of proteins at relatively large scales, involving tens of nanometres and with full atomistic resolution. Importantly, the development of this large model based directly on X-ray diffraction data is an advance over earlier smaller models in which the stutter was placed closer to the boundary of the filament (Ackbarow and Buehler 2009). The vimentin coil2 segment used here is obtained by combining two coiled-coil segments obtained directly from Protein Data Bank (PDB). The N-terminal half of coil2 is obtained from vimentin segment that includes residues 264–335 (obtained from the PDB with identification (ID) 3KLT Nicolet et al. 2010). The C-terminal half of coil2 is obtained from the vimentin segment that includes residues 328–406 (given by PDB with ID 1GK4 Strelkov et al. 2002). The two segments are aligned to form an integrated coiled coil because both fragments include an overlapped region of residues 328–335 as shown in Figure 1(B). We use visual molecular dynamics (Humphrey et al. 1996) to find the best fit between the backbone atoms of the overlapped region of the two segments corresponding to the overlapped part. The best fit provides a minimum root mean square deviation of 0.68 Å between the C-alpha atoms within the overlapped region of the two segments as shown in Figure 1(C). We find that this result is better than directly superimposing the two segments (of 1.3 Å), and both of them indicate that the two segments match well with each other.

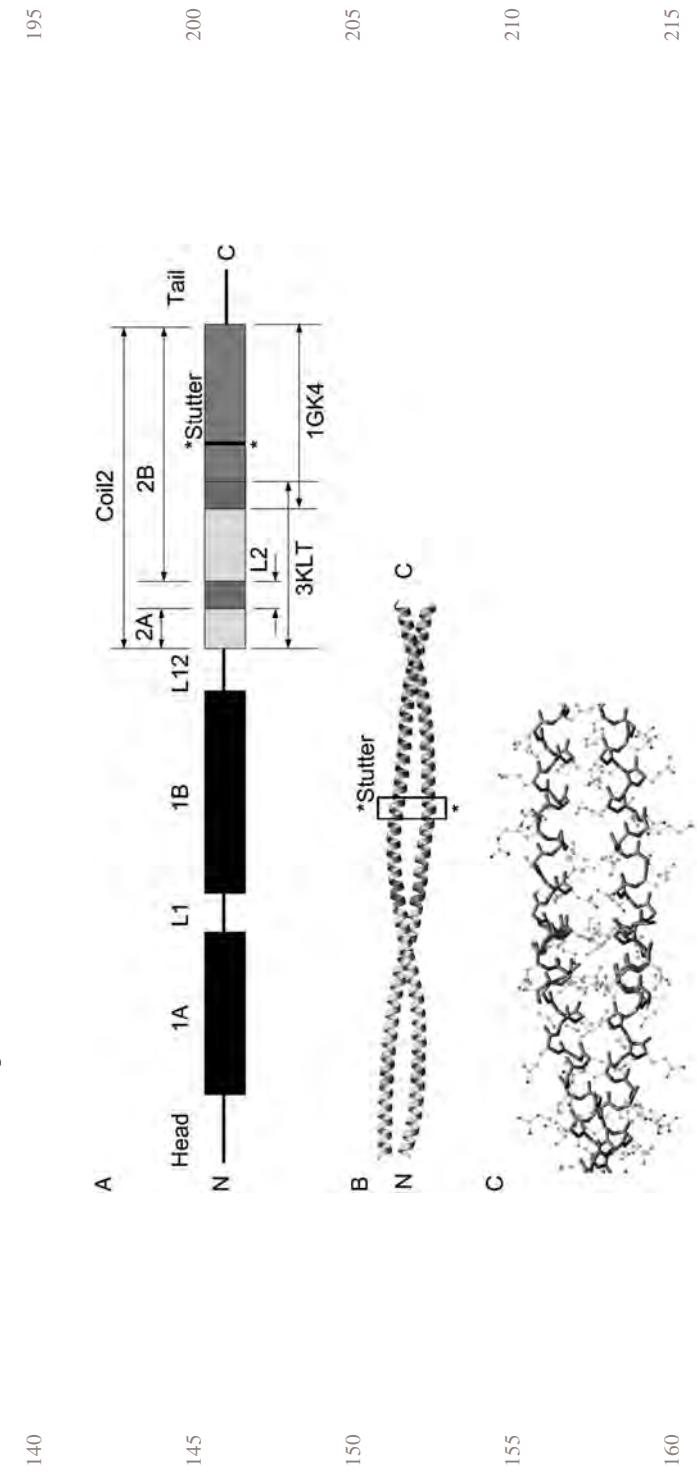


Figure 1. Schematic view of human vimentin dimer and the used segments in this study, with the location of the stutter region pointed out by the \* symbol. Panel (A): schematic view of the human vimentin IF dimer. Panel (B): cartoon view of the entire coil2 domain in the vimentin dimer. Note that in Panels (A) and (B) the red regions correspond to the overlapped region (residues 328–335) as we use to integrate the two segments from PDB. Panel (C): detailed view at the overlapped region of the two segments after alignment.

Molecular simulation approach

We use a standard explicit water molecular dynamics approach to first carefully equilibrate and then mechanically deform the coil2 of vimentin. The structure is 21 nm long in its initial configuration. All molecular simulations are carried out using the CHARMM force field (MacKerell et al. 1998) as implemented in NAMD (Nelson et al. 1996). We use a 2 fs time step at a constant temperature of  $T = 300\text{K}$  within a  $NPT$  ensemble. A pressure of  $P = 1\text{ atm}$  is used. We start the simulation protocol by equilibrating the molecular structure for 20 ms, where Figure 2 shows the resulting structure, including the constituent domains and the hydrophobic regions. We find that the 20 ms equilibration, limited due to computational constraints, gives a well-equilibrated structure with good geometry convergence. The molecular configuration obtained after equilibration is subsequently deformed using steered molecular dynamics (SMD) (Lu et al. 1998). In the simulation to deform the structure using SMD, the C-alpha atoms at the head of the structure for both constituting chains are fixed whereas the C-alpha atoms at the tail of the structure are pulled up to a strain of 30%. The SMD spring constant used is  $k = 10\text{kcal mol}^{-1}\text{\AA}^{-1}$  in agreement with earlier studies. Varied pulling rates of 0.5, 1.0, 5 and 10 m/s are considered to present the validity of the approach and as a comparison with the force levels observed in earlier studies.

## *Free-energy calculations*

In addition to SMD simulations, free-energy calculations are carried out using well-tempered metadynamics (Barducci et al. 2008) implemented in the PLUMED (Bonomi et al. 2009) plugin for NAMD, also at  $T = 300\text{ K}$ . Tests are performed for two different sections in the 2B domain in order to identify their free energy in the folded and unfolded state. The regions studied include the stutter and a set of residues neighbouring the stutter region that is known to form a stable alpha helix. The collective variable in metadynamics is chosen to be the distance between the C-alpha atoms of residues 349–353 for the stutter, and 363–367 for the other section. The tests are run for an initial configuration of perfectly folded alpha helix and also for the configuration obtained after equilibration for 20 ns. The PLUMED runs here are 5 ns long and the convergence of the tests is confirmed.

## Results and discussion

We begin the analysis with a presentation of the structure after equilibration as shown in Figure 2. The structure maintains the alpha-helical secondary structure in most of the system. However, we observe that the stutter region unfolds locally even without any applied force. This indicates a reduced stability of the alpha-helical secondary structure in this region and clearly reveals the formation of a structural defect at the site of the stutter. This result suggests that the stutter may not only be a region in which alpha helices run in parallel as proposed earlier, but that the stutter region may actually represent a complete, albeit local, loss of the alpha-helical structure.

In order to gain better insight into the unfolding observed during equilibration, free-energy calculations are performed that allow us to sample a much longer time-scale than accessible to standard molecular dynamics. Figure 3 shows the free-energy profile of the stutter region compared with a reference case that represents a neighbouring alpha-helix region in which no instability is expected. This allows us to investigate the free-energy levels for an ‘unfolded’ and ‘folded’ state. In the ‘folded’ state, the end-to-end distance for one turn of alpha helix (corresponding to four residues) is around 0.6 nm. In the ‘unfolded’ state, this distance fluctuates around 1.2 nm, as shown in Figure 3(C). For each case (stutter and reference), two different test

Figure 2. Panel (A): in the simulation unit locuses on deforming the structure under uniaxial strain, the C-alpha atoms at the head of the molecule for both the constituent chains are fixed, whereas the C-alpha atoms at the tail of the structure are pulled. Panel (B): the equilibrated structure with the constituent segments marked (2A, L2 and 2B) with an inset showing the detailed view of the stutter region. The structure was equilibrated for 20 ns in explicit water solvent and the 2A region is found to be in a relatively rigid and parallel configuration, whereas the 2B region is not aligned with any orthogonal axis and has gone through local deformations, especially at the vicinity of the stutter region which locally unfolds even without any applied force.

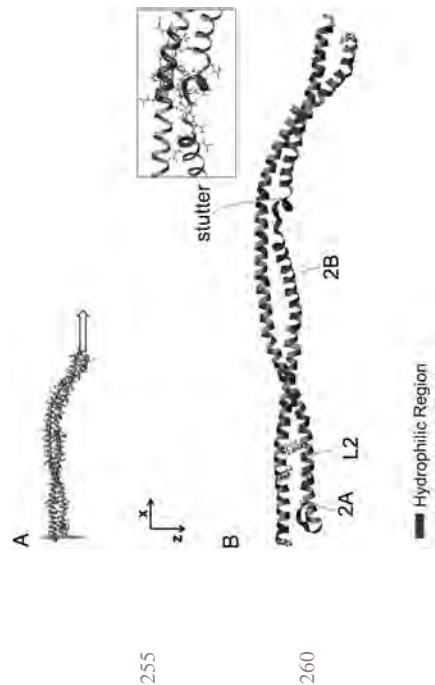


Figure 2. Panel (A): in the simulation that focuses on deforming the structure under uniaxial strain, the C-alpha atoms at the head of the molecule for both the constituent chains are fixed, whereas the C-alpha atoms at the tail of the structure are pulled. Panel (B): the equilibrated structure with the constituent segments marked (2A, L2 and 2B) with an inset showing the detailed view of the stutter region. The structure was equilibrated for 20 ns in explicit water solvent and the 2A region is found to be in a relatively rigid and parallel configuration, whereas the 2B region is not aligned with any orthogonal axis and has gone through local deformations, especially at the vicinity of the stutter region which locally unfolds even without any applied force.

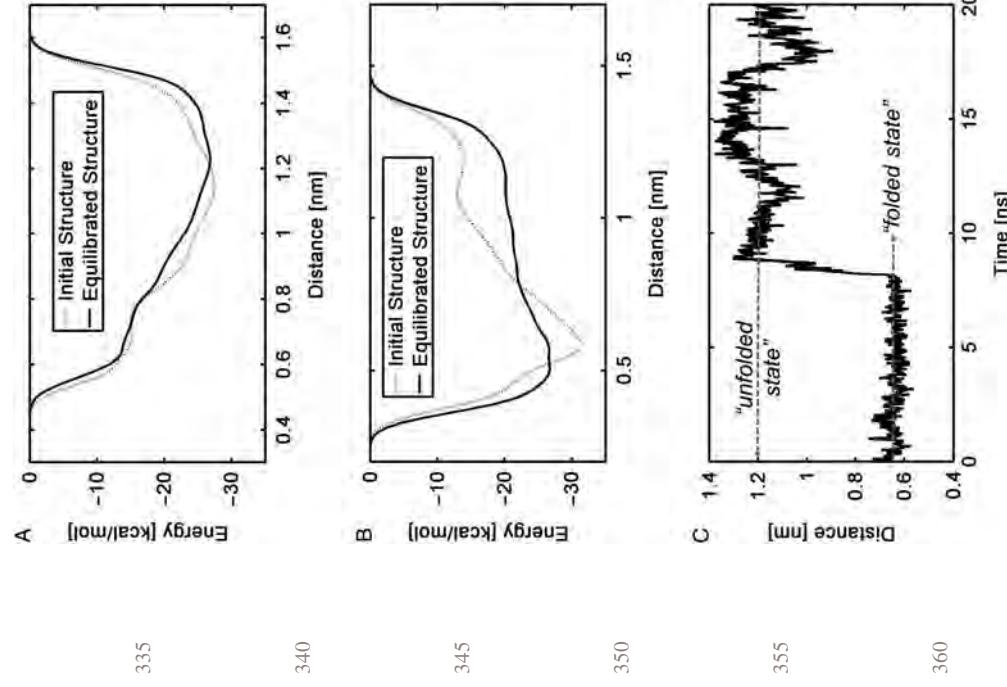


Figure 3. Metadynamics analysis of the free-energy landscape of the stutter (A) and a reference region with perfect alpha-helical coiled coils (B). Panel (A) shows the free-energy profile for the stutter region. The x-axis of the plot denotes the distance of one turn (residues 349–353) where the stutter exists. The free-energy landscape shows a minimum energy configuration at the ‘unfolded’ state, suggesting that it is more stable. Panel (B) shows the free-energy profile for the distance of one turn that constitutes residues 363–367 in the reference region. The results reveal that the minimum of the free-energy profile corresponds to the folded state for the alpha helix. Panel (C) shows the 20 ns equilibration result revealing the distances associated with the ‘folded’ (around 0.6 nm distance) and unfolded (around 1.2 nm distance) states, respectively.

3(B)), and the results show a minimum energy level at the ‘folded’ state, in contrast to the stutter case.

In addition to the different minimum free-energy locations for each case, this analysis also portrays the difference in the energy barrier to go from a ‘folded’ state to an ‘unfolded’ state between two different sets of residues. The energy level difference to go from a ‘folded’ state to an ‘unfolded’ state being  $\sim 7$  kcal/mol for the stutter region (Figure 3(A)), and that for the reference case

is shown to be  $\sim 11$  kcal/mol. (Figure 3(B)); for the analysis performed starting from an equilibrated state. We note that when the analysis is carried out for the initially ‘unfolded’ state, we observe that the ‘folded’ state that the stutter visits during the metadynamics run is a 390 slightly different folded configuration of a perfect alpha helix. The well-tempered metadynamics analysis presented here provides a powerful tool in determining the free-energy landscape associated with different states of the lengths of the regions of vimentin coil2, and confirms 395 that the stutter region readily unfolds during equilibration because its free energy in the unfolded state is lower.

We now proceed with a presentation of a detailed 400 nanomechanical analysis of the coil2 domain. Figure 4 shows force-displacement curve for the structure. In agreement with earlier studies (Ackbarow and Buehler 2009; Bertaud et al. 2010), we find that the curve has two 405 distinct regions, where the force increases linearly until the structure is aligned with the pulling axis and followed by a small regime of axial stretching. The alignment of the structure with the pulling axis occurs at approximately 2 nm displacement, which is equivalent to about 10% strain for the 21 nm long structure (it is noted that the structure is not aligned with any orthogonal axis system 410 after equilibration as seen in Figure 2).

After the initial almost linear-elastic region of the force-displacement curve in Figure 4 the angular point is reached, where a further increase in the prescribed displacement does not result in an increase in the force. 415 This is due to the subsequent unfolding events that involve rupture of H-bonds and as a result, alpha-helical turns that define the structure of the alpha helices. We observe that the rupture of alpha-helical turns begins at the stutter and in the vicinity of the stutter region, as well as at the linker (L2) part at a displacement of about 4 nm (this implies a 420 strain level of about 20% for the 21 nm long structure).

This observation suggests that the initial defect formed at the stutter serves as a seed for the initiation of molecular 425 unfolding under applied force. The data shown in Figure 4(A) suggest an average unfolding force of about 350 pN. For the pulling rate of 1 m/s used here, this unfolding force is in accordance with the results from the previous studies (Bertaud et al. 2010). Figure 4(B) depicts the results of the same analysis performed over a range of different pulling 430 rates. It is shown that the trend of initial linear-elastic region leads to a plateau over which the unfolding of different sections of the protein happens over a range of different pulling rates (0.5, 1, 5 and 10 m/s). However, the plateau force increases with increasing pulling rate as 435 studied and shown in earlier work for protein unfolding simulations in explicit water (Bertaud et al. 2010).

Figure 5 displays a detailed analysis of chemical and structural changes during deformation, based on the percentage of change of the H-bonds in the backbone of 440 the structure as well as the secondary structure transition

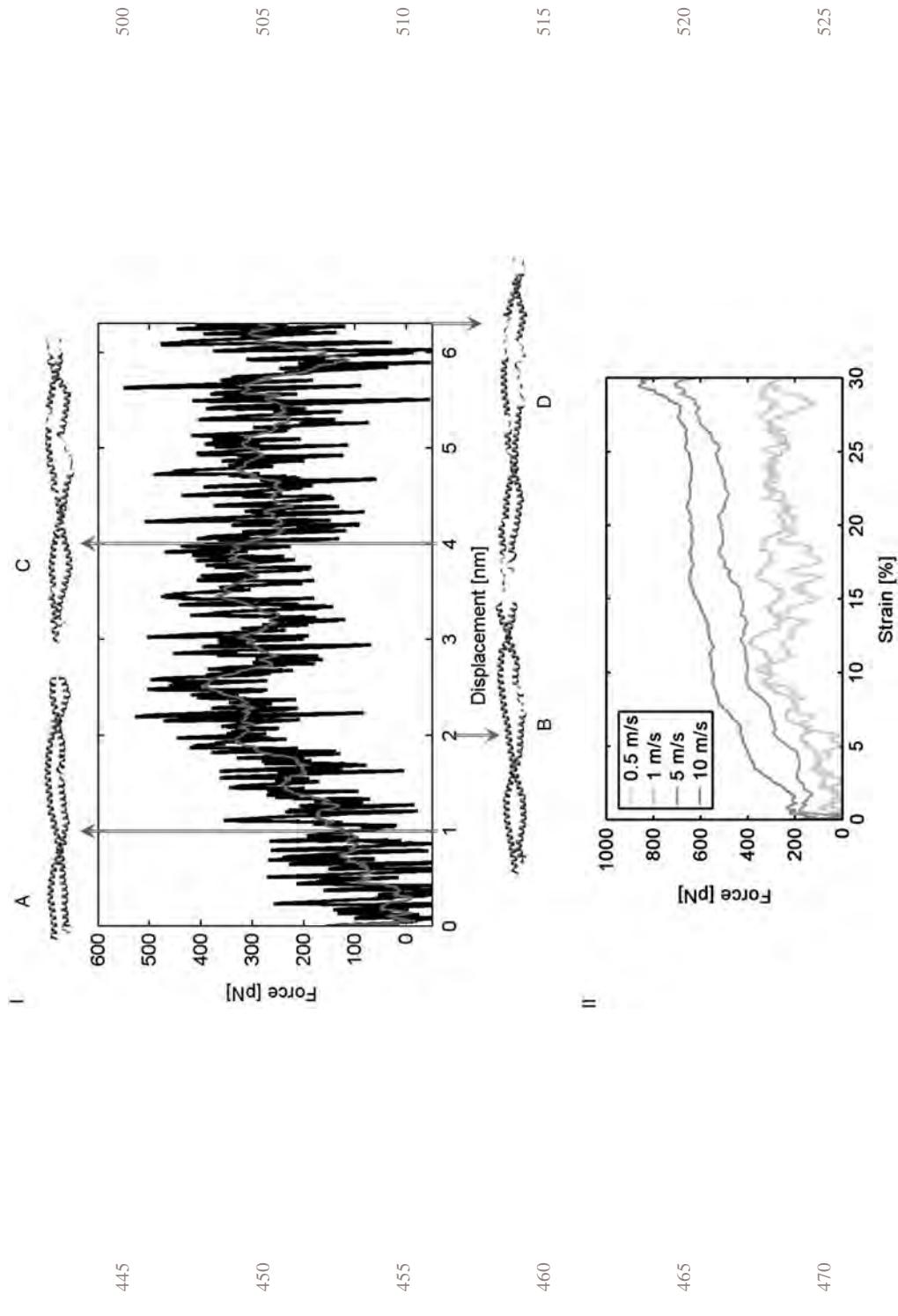


Figure 4. (I) Force-displacement curve of the coil2 structure, showing (A) at a displacement of 1 nm, the structure is still aligned with the pulling axis and the stutter region is readily unfolded from the initial alpha helix form, (B) further increase in displacement results in alignment with the pulling axis as well as further unfolding of the stutter region which also defines the unfolding force (about 350 pN) which is constant throughout the subsequent unfolding events, (C) at a displacement of about 4 nm, the linker region starts unfolding as well as the 2B region, (D) further increase in displacement results in an unfolded 2B region and an unstable linker region whereas the 2A regions retain its form (i.e. alpha-helix configuration). (II) Force-displacement plots of tensile test simulations for four different pulling velocities (0.5, 1, 5 and 10 m/s). The plateau region depicting the consecutive unfolding events is at a significantly higher force level for higher pulling speeds such as 5 or 10 m/s.

due to the unfolding of the alpha-helix domains. Figure 5(A) shows the plot where the H-bonds breakage events are monitored in the duration of equilibration of the structure (20 ns). The negative change in the percentage of the H-bonds indicates the unfolding of the alpha helix along the backbone of the structure. A much higher ratio (about 25%) of H-bond rupture is observed as expected when the deformation is applied (Figure 5(B)). Here, again, the trend of the protein unfolding is dictated by the trend in the 2B region of the protein. This assures us that there are no significant unfolding events that occur in the 2A section of the structure up to strains of 30%. Figure 5(C) marks the increase in the random coil structure due to

the unfolding of the alpha-helix domains with the applied strain, in agreement with the data provided in Figure 4.

### Conclusion

We presented an explicit water molecular simulation analysis of the deformation of a 21 nm long vimentin coil2 domain and observed several striking results. Most importantly, we found that the stutter region readily unfolds even under the absence of any externally applied force, suggesting that the stutter is not very stable in the alpha-helical state. We confirmed this hypothesis using well-tempered metadynamics, which enabled us to obtain the free-energy configuration for an ‘unfolded’ and a

aligned with the pulling axis and the stutter region is readily unfolded from the initial alpha helix form, (B) further increase in displacement results in alignment with the pulling axis as well as further unfolding of the stutter region which also defines the unfolding force (about 350 pN) which is constant throughout the subsequent unfolding events, (C) at a displacement of about 4 nm, the linker region starts unfolding as well as the 2B region, (D) further increase in displacement results in an unfolded 2B region and an unstable linker region whereas the 2A regions retain its form (i.e. alpha-helix configuration). (II) Force-displacement plots of tensile test simulations for four different pulling velocities (0.5, 1, 5 and 10 m/s). The plateau region depicting the consecutive unfolding events is at a significantly higher force level for

485

490

495

530

545

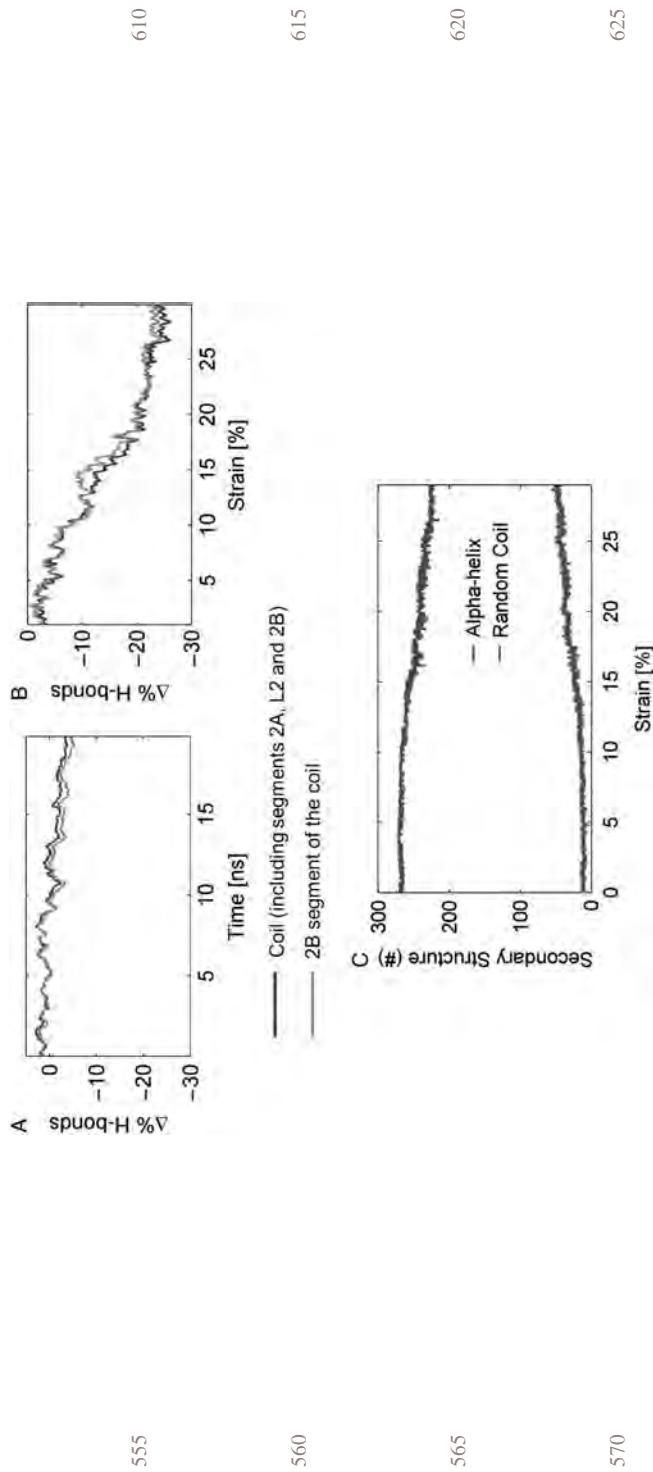


Figure 5. Percentage change in the number of backbone H-bonds in the duration of (A) equilibration (20 ns) and (B) deformation (plotted over the macroscopic strain of 30%). In panel (A), the breaking of the H-bonds is observed during equilibration (amounting to about 5%), the trend in the decrease of the percentage change in the number of H-bonds for the structure (the blue curve) is the same as that of segment 2B (red curve). This implies that the breaking of the H-bonds happens in segment 2B. Panel (B) shows that there is a nonlinear decrease (up to about 25% change in the amount of the H-bonds) in the breakage of the H-bonds during deformation (the breakage is slower as the deformation reaches about 25% strain). Here, we can observe that the trend of H-bonds breakage is governed by segment 2B which experiences consecutive unfolding events. Note that in panel (B), the reference configuration is the one at the end of equilibration. Panel (C) shows the secondary structure content as a function of strain for the entire coil2 (2A, L2 and 2B). As the strain increases, the number of alpha-helix domains decreases whereas the number of random coil domains (which do not correspond to a specific secondary structure) increases. This point at approximately 15% strain marks the occurrence of the unfolding events and the complete breakdown of the initial alpha-helical geometry.

‘folded’ state for the stutter region compared with a regular alpha-helical region that serves as a reference case. The analysis of these results highlighted that the stutter region is more stable in the ‘unfolded’ state, in agreement with the observations made in regular molecular dynamics simulations (Figure 3). By comparison, the analysis of the reference case confirms that the ‘folded’ state is more stable.

The thinner cross section of the stutter region represents a weak point as it forms a location with a reduced local bending stiffness. This could affect the persistence length along the strands which then affects the mobility of the structure during equilibration. Possible biological implications of this defined local unfolding may relate to signalling processes, where large changes in the molecular conformation may induce biochemical signalling cascades to indicate mechanical strain in varied cell mechanical states. A detailed analysis of these issues and further experimental validation of the predictions made here are left for future work.

Another interesting outcome of the simulation results presented here is the relative rigidity of the 2A region, as anticipated earlier (Qin et al. 2009). The 2A region retains

its alpha-helix form even up to a strain of 30% (at a displacement of about 6 nm), which shows the relatively stiff nature of this region in the dimer. The work presented here is a part of a larger effort to elucidate structure–property relationships in hierarchical materials, an effort defined as materionics that could eventually result in an improved understanding of material factors in physiological conditions and disease (Cranford and Buehler 2010).

#### Acknowledgements

This work was supported by DOD-MURI, AFOSR and a PECASE grant. Additional support from NSF is acknowledged.

#### References

- Ackbarow T, Buehler MJ. 2009. Molecular mechanics of stutter defects in vimentin intermediate filaments. *Exp Mech.* 49(1):79–89.
- Ackbarow T, Sen D, Thaulow C, Buehler MJ. 2009. Alpha-helical protein networks are self protective and flaw tolerant. *PLoS ONE.* 4(6):e6015.
- Alberts B. 2002. Molecular biology of the cell. New York: Taylor and Francis.

Barducci A, Bussi G, Parrinello M. 2008. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett*. 100(2):020603.

Bertaud J, Hester J, Jimenez DD, Buehler MJ. 2010. Energy landscape, structure and rate effects on strength properties of alpha-helical proteins. *J Phys: Condens Matter*. 22:035102.

Bonomi M, Branduardi D, Bussi G, Camilloni C, Provasi D, Raiteri P, Donadio D, Marinelli F, Pietrucci F, Broglia RA, Parrinello M. 2009. PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun*. 180(10):1961–1972.

**Q4** Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A. 2001. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. *Nat Genet*. 27(1):117–120.

Brown JH, Cohen C, Parry DAD. 1996. Heptad breaks in alpha-helical coiled coils: stutters and stammers. *Proteins-Struct Funct Genet*. 26(2):134–145.

Buehler MJ, Ackbarow T. 2008. Nanomechanical strength mechanisms of hierarchical biological materials and tissues. *Comput Methods Biomech Biomed Engin*. 11(6):595–607.

Buehler MJ, Yung YC. 2009. Deformation and failure of protein materials in physiologically extreme conditions and disease. *Nat Mater*. 8(3):175–188.

Canford SW, Buehler MJ. 2010. Materomics: biological protein materials, from nano to macro. *Nanotechnol Sci Appl*. 3:127–148.

Dahl KN, Kahn SM, Wilson KL, Discher DE. 2004. The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. *J Cell Sci*. 117(20):4779–4786.

Fudge D, Russell D, Beriault D, Moore W, Lane EB, Vogl AW. 2008. The intermediate filament network in cultured human keratinocytes is remarkably extensible and resilient. *PLoS ONE*. 3(6):e23227.

Herrmann H, Aebi U. 2004. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. *Annu Rev Biochem*. 73:749–789.

Herrmann H, Bar H, Kreplak L, Strelkov SV, Aebi U. 2007. Intermediate filaments: from cell architecture to nanomechanics. *Nat Rev Mol Cell Biol*. 8(7):562–573.

Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J Mol Graph*. 14(1):33–38.

Hutchison CJ. 2002. Lamins: building blocks or regulators of gene expression? *Nat Rev Mol Cell Biol*. 3(11):848–858.

Kreplak L, Herrmann H, Aebi U. 2008. Tensile properties of single desmin intermediate filaments. *Biophys J*. 94(7):2790–2799.

Lewis MK, Nahirney PC, Chen V, Adhikari BB, Wright J, Reedy MK, Bass AH, Wang K. 2003. Concentric intermediate filament lattice links to specialized Z-band junctional complexes in sonic muscle fibers of the type I male midshipman fish. *J Struct Biol*. 143(1):56–71.

Lu H, Isralewitz B, Krammer A, Vogel V, Schulten K. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys J*. 75(2):662–671.

MacCallum JL, Moghaddam MS, Chan HS, Tielemans DP. 2007. Hydrophobic association of alpha-helices, steric dewetting, and enthalpic barriers to protein folding. *Proc Natl Acad Sci USA*. 104(15):6206–6210.

Mackellar AD et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*. 102(18):3586–3616.

**Q7** Nelson MT, Humphrey W, Gursoy A, Dalke A, Kalé LV, Skeel RD, Schulten K. 1996. NAMD: a parallel, object oriented molecular dynamics program. *Int J Supercomput Appl High Perform Comput*. 10(4):251–268.

Nicolet S, Herrmann H, Aebi U, Strelkov SV. 2010. Atomic structure of vimentin coil 2. *J Struct Biol*. 170(2):369–376.

Omary MB, Coulombe PA, McLean WH. 2004. Intermediate filament proteins and their associated diseases. *N Engl J Med*. 351(20):2087–2100.

Parry DAD, Strelkov SV, Burkhardt P, Aebi U, Herrmann H. 2007. Towards a molecular description of intermediate filament structure and assembly. *Exp Cell Res*. 313(10):2204–2216.

**Q5** Qin Z, Kreplak L, Buehler MJ. 2009. Nanomechanical properties of vimentin intermediate filament dimers. *Nanotechnology*. 20(42):425101.

Rose A, Meier I. 2004. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. *Cell Mol Life Sci*. 61(16):1996–2009.

Strelkov S. 2002. Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. *EMBO J*. 21(6):1255–1266.

Strelkov SV, Herrmann H, Geisler N, Lustig A, Ivaninskii S, Zimbelmann R, Burkhardt P, Aebi U. 2001. Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. *J Mol Biol*. 306(4):773–781.

Strelkov SV, Herrmann H, Geisler N, Wedig T, Zimbelmann R, Aebi U, Burkhardt P. 2002. Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. *EMBO J*. 21(6):1255–1266.

Strelkov SV, Herrmann H, Aebi U. 2003. Molecular architecture of intermediate filaments. *Bioessays*. 25(3):243–251.

Wang N, Butler JP, Ingber DE. 1993. Mechanotransduction across the cell surface and through the cytoskeleton. *Science*. 260(5111):1124–1127.

Wang N, Stamenovic D. 2002. Mechanics of vimentin intermediate filaments. *J Muscle Res Cell Motil*. 23(5–6):535–540.

Wilson KL, Zastrow MS, Lee KK. 2001. Lamins and disease: insights into nuclear infrastructure. *Cell*. 104(5):647–650.